University of Alberta

Design of an RF Wideband Low-Noise Amplifier Using 0.35μm CMOS Technology

by

Yan Tang ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Master of Science**.

Department of Electrical and Computer Engineering

Edmonton, Alberta
Spring 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-69768-1

Canada

University of Alberta

Library Release Form

Name of Author: Yan Tang

Title of Thesis: Design of an RF Wideband Low-Noise Amplifier Using 0.35μm CMOS Technology

Degree: Master of Science

Year this Degree Granted: 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Tan Yan

Yan Tang
230 Burton Road
Edmonton, Alberta
Canada T6R 1N9

Date: Dec. 17, 2001

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Design of an RF Wideband Low-Noise Amplifier Using 0.35μm CMOS Technology** submitted by Yan Tang in partial fulfillment of the requirements for the degree of **Master of Science**.

Dr. Igor M. Filanovsky

Dr. Al Meldrum

Dr. James N. Mcmullin

Date: Dec. 14, 2001

To My Parents

Abstract

This thesis investigates the characteristics and details the design process of a full-CMOS wideband low-noise amplifier (LNA) for application in a wireless RF receiver front-end. Several different LNA architectures and wideband amplification methods are surveyed. A new technique is investigated using a bridged T-Coil which can be realized by a pair of magnetically coupled inductors (i.e., a transformer) for bandwidth extension . The spiral inductors and transformer have been implemented. The prototype has been designed and fabricated using a 0.35μm CMOS technology.

This wideband LNA achieves 11.64 dB forward gain and the -3dB bandwidth ranges from 0.3 MHz to 642 MHz. The reverse isolation is less than -32 dB. It consumes 14.42 mW from a 2 V supply.

Acknowledgements

I would like to take this opportunity to express my deep gratitude to my supervisor, Dr. Igor M. Filanovsky, for his guidance, encouragement and insight, and also for his patience in seeing this work through. This project has been an incredibly rewarding and challenging experience from beginning to end and I am grateful to have had this experience.

Many thanks are given to the Canadian Microelectronics Corporation for the help, technical support and the fabrication of chips, and to TRLabs in Edmonton for providing lab testing environment.

One of the greatest blessings in my life has been the influence of my family. Consequently, I wish to express my deepest appreciation to my parents and my brother Wei, for their love, support, and what they have done for me. They have been extremely important not only in making me who I am, but also in helping me through the highs and lows that have accompanied my university life.

Contents

1 Introduction **1**
- 1.1 Architecture of an RF System . . . 1
- 1.2 Low Noise Amplifier Fundamentals . . . 2
  - 1.2.1 Noise Figure . . . 3
  - 1.2.2 Gain . . . 5
  - 1.2.3 Bandwidth . . . 5
  - 1.2.4 Linearity . . . 7
  - 1.2.5 Input and Output Matching . . . 10
- 1.3 CMOS Fundamentals . . . 10
  - 1.3.1 Basic Equations . . . 10
  - 1.3.2 Noise Sources in CMOS and Bipolar Devices . . . 12
- 1.4 Thesis Organization . . . 12

2 Wideband Amplification and Matching **14**
- 2.1 Introduction . . . 14
- 2.2 Wideband Amplification Methods . . . 14
  - 2.2.1 Local Feedback Configuration . . . 15
  - 2.2.2 Overall Feedback Configuration . . . 17
  - 2.2.3 Compound Transistors . . . 18
  - 2.2.4 Inductive Peaking . . . 20
- 2.3 Common LNA Architectures . . . 24
  - 2.3.1 Common-Gate Architecture . . . 24
  - 2.3.2 Resistive Termination Architecture . . . 25
  - 2.3.3 Shunt-Series Feedback Architecture . . . 26
  - 2.3.4 Inductive Degeneration Architecture . . . 27
- 2.4 Conclusion . . . 29

3 Inductor and Transformer Design **32**
- 3.1 Introduction . . . 32
- 3.2 Inductor Requirements . . . 33
- 3.3 Inductor and Transformer Modeling . . . 33
  - 3.3.1 Equivalent Circuits of Spiral Inductors . . . 33
  - 3.3.2 Equivalent Circuit of Coupled Inductors . . . 34
  - 3.3.3 Comparison of Different Transformer Layouts . . . 35
- 3.4 Design Guidelines . . . 38

3.5 Inductor and Transformer Performance 40
3.5.1 Summary of the Measurements 40
3.5.2 Advice for Laying Out and Testing Inductors and Transformers 41

4 Design Goals and Approach **47**
4.1 Introduction 47
4.2 Design Goals 47
4.3 Approach 48
4.3.1 Process Choice 48
4.3.2 Multi-Stage Design 50

5 Implementation **51**
5.1 Circuit Design 51
5.1.1 Transistor Biasing 51
5.1.2 Input Stage 53
5.1.3 Output Stage 55
5.1.4 Theoretical Calculations 59
5.2 Wideband LNA Layout 63
5.2.1 Active Device Layout 63
5.2.2 Capacitor Layout 65
5.2.3 Resistor Layout 66
5.2.4 Final Circuit Schematic and Layout 68
5.3 Simulation Results 69

6 Experimental Results **76**
6.1 Overview 76
6.2 Test Equipment 77
6.3 Chip Testing 77

7 Conclusions **87**
7.1 Thesis Summary 87
7.2 Improvements and Future Work 88

Bibliography **90**

A Appendix **94**

B Appendix **113**

List of Figures

1.1 Block diagram of a common RF transceiver: (a) transmitter, (b) receiver. 2
1.2 Noise figure calculation for an amplifier. 3
1.3 Graphical representation of 1-dB compression point, IIP3. 9

2.1 Transconductance configuration. 15
2.2 Transimpedance configuration. 16
2.3 Voltage amplifier. 17
2.4 Current amplifier. 18
2.5 Two-stage voltage amplifier with global feedback. 19
2.6 Two-stage current amplifier with global feedback. 20
2.7 Common-drain-common-source configuration. 21
2.8 Cascode configuration. 22
2.9 Basic CS-voltage amplifier configuration. 23
2.10 Shunt-peaking amplifier. 24
2.11 Series-peaking amplifier. 25
2.12 Shunt-series-peaking amplifiers. 26
2.13 Common LNA architectures: (a) common-gate stage, (b) resistive termination, (c) shunt-series feedback, (d) inductive degeneration. 30
2.14 The equivalent circuit of the inductively degenerated common-source amplifier. 31
2.15 Schematic diagram used for effective transconductance calculations. . 31

3.1 A spiral inductor layout. 34
3.2 Traditional spiral inductor model. 35
3.3 Modified spiral inductor model. 36
3.4 Layout of two coupled spiral structures. 36
3.5 Equivalent circuit model for two coupled spirals. 37
3.6 Transformer layouts I: (a) Tapped, (b) Interleaved, (c) Stacked with top spiral overlapping the bottom one, (d) Stacked with top spiral laterally shifted, (e) Stacked with top and bottom spirals diagonally shifted. 44
3.7 Transformer Layouts II. 45
3.8 Test Layout of Inductors and Transformer. 45
3.9 Equivalent T-circuit model for a transformer. 46

5.1 Implementation of voltage dividers in CMOS. 52

5.2 Modification of the resistor-MOSFET voltage divider. 53
5.3 Input stage of the LNA. 54
5.4 Output stage of the LNA. 56
5.5 (a) Coupled inductors forming a T-coil, (b) Equivalent circuit with mutual inductance M. 57
5.6 (a) Bridged T-coil with C_L load, (b) Equivalent circuit. 71
5.7 Biasing circuit for the cascoded input stage of the LNA. 72
5.8 Connecting four transistors in parallel to realize a single large transistor: (a) the layout, (b) the schematic drawn in the same relative positions as the layout, and (c) the circuit redrawn to make the parallel transistors more obvious. 73
5.9 Plan view of a typical poly-poly capacitor. 74
5.10 A typical layout for an integrated poly-silicon resistor. 74
5.11 Complete LNA1. 75
5.12 Complete LNA2. 75

6.1 General setup for electronic measurements. 77
6.2 Frequency response experimental setup. 81
6.3 Wideband frequency response calibration with 10 dB attenuation. . . 82
6.4 S_{11} vs. frequency for LNA1. 83
6.5 S_{21} vs. frequency for LNA1. 84
6.6 S_{12} vs. frequency for LNA1. 85
6.7 S_{22} vs. frequency for LNA1. 86

A.1 Layout of the 360μm/0.35μm transistor. 95
A.2 Layout of the 90fF poly-poly capacitor. 96
A.3 Layout of the 200Ω poly-silicon resistor. 97
A.4 Final schematic of LNA1 with components values. 98
A.5 Final schematic of LNA1 with DC operating points. 99
A.6 Final layout of LNA1. 100
A.7 Part A in Figure A.6 in detail. 101
A.8 Part B in Figure A.6 in detail. 102
A.9 Final schematic of LNA2 with components values. 103
A.10 Final schematic of LNA2 with DC operating points. 104
A.11 Final layout of LNA2. 105
A.12 Part A in Figure A.11 in detail. 106
A.13 Part B in Figure A.11 in detail. 107
A.14 (a) V_{out} vs. frequency for LNA1, (b) Z_{in} vs. frequency for LNA1. . . 108
A.15 Z_{out} vs. frequency for LNA1. 109
A.16 (a) V_{out} vs. frequency for LNA2, (b) Z_{in} vs. frequency for LNA2. . . 110
A.17 Z_{out} vs. frequency for LNA2. 111
A.18 Pad layout used throughout the project. 112

B.1 Die photo for LNA1. 114
B.2 Die photo for the inductor with inductance of 10nH. 115

B.3 Die photo for the stacked transformer (P=9nH; S=9nH; k=1/3). . . . 116

List of Tables

3.1 Comparison between different transformer layouts 37
3.2 0.35μm CMOS Process Parameters 40
3.3 Layout Geometry of Inductors and Transformer 41
3.4 Simulation results for the inductors 41
3.5 Simulation results for the transformer 41
3.6 Measurement results for inductor 1 42
3.7 Measurement results for inductor 2 42
3.8 Measurement results for transformer primary winding 42
3.9 Measurement results for transformer secondary winding 43

5.1 Design circuit parameters 69
5.2 Summary of simulation results for two LNAs ($V_{cc} = 2V$; $V_{in} = 100\mu V$) 70

6.1 Summary of LNA1 measurements 79

List of Symbols

F:	noise factor
P_{si}:	input signal power
P_{ni}:	input noise power
P_{so}:	output signal power
P_{no}:	output noise power
G_a:	frequency-dependent available power gain
P_{ne}:	output noise power due to internal noise sources
$\tilde{v}_n^2$:	amplifier input-referred noise-voltage generator
$\tilde{i}_n^2$:	amplifier input-referred noise-current generator
k:	Boltzmann's constant
T:	absolute temperature in Kelvin
R_s:	source resistance
Δf:	bandwidth in hertz over which the noise is measured
ρ:	correlation coefficient between $\tilde{v}_n^2$ and $\tilde{i}_n^2$
V_n:	root-mean-square noise voltage
I_n:	root-mean-square noise current
G:	power gain
A_v:	voltage gain
v_o:	instantaneous value of the output voltage
f_H/f_L:	half-power (or -3 dB) frequencies
$H(\jmath\omega)$:	transfer function
Γ:	reflection coefficient
V_{gs}:	gate-to-source voltage
V_{ds}:	drain-to-source voltage
V_t:	threshold voltage
I_D:	drain current of the saturated NMOS transistor
μ_n:	mobility
C_{ox}:	gate-oxide capacitance per unit area
W:	width of the channel
L:	length of the channel
g_m:	transconductance of the device
ω_T:	unity current-gain frequency
C_{gs}:	capacitance between the gate and the source
g_{d0}:	zero-bias drain conductance of the device
γ:	bias-dependent coefficient
τ:	time constant

t_r:	10 to 90 % risetime in response to a unit step input
ω_{-3dB}:	-3 dB bandwidth in terms of angular frequency
ω_0:	midband resonant frequency of operation
G_m:	effective transconductance
Q:	quality factor
$k_{1,2}$:	mutual coupling coefficient between two inductors
M:	mutual inductance
l:	metal line length
w:	metal line width
s:	metal line spacing
N:	number of turns
R_{sh}:	metal sheet resistance
C_{sub}:	parasitic capacitance from metal to substrate
t:	thickness of metal or layer
ϵ:	relative permittivity
f_{max}:	maximum oscillation frequency
t_{ox}:	gate-oxide thickness
ε_{ox}:	dielectric constant of the gate oxide
R_g:	gate resistance
$R_{\square}$:	poly-silicon sheet resistance
W_u:	unit finger width

List of Acronyms

ADC:	Analog-to-Digital Converter
AMPS:	Advanced Mobile Phone Service
ASITIC:	Analysis and Simulation of Inductors and Transformers in Integrated Circuits
BiCMOS:	Bipolar-Complementary-Metal-Oxide-Semiconductor
BW:	Bandwidth
CMOS:	Complementary-Metal-Oxide-Semiconductor
DAC:	Digital-to-Analog Converter
DECT:	Digital European Cordless Telephone
DR:	Dynamic Range
DUT:	Device Under Test
FET:	Field-Effect-Transistor
GaAs:	Gallium Arsenide
GBP:	Gain-Bandwidth Product
GMD:	Geometric Mean Distance
GS:	Ground-Signal
GSG:	Ground-Signal-Ground
GSM:	Group Special Mobile
IIP3:	Input Third-Order Intercept Point
IMD:	Intermodulation Distortion
IP3:	Third-Order Intercept
ISS:	Impedance-Standard-Substrate
LAN:	Local Area Network
LNA:	Low Noise Amplifier
LRL:	Line-Reflect-Line
LRM:	Line-Reflect-Match
MFA:	Maximally Flat Amplitude
MFED:	Maximally Flat Group Delay
NF:	Noise Figure
NFET:	N-Channel Field-Effect-Transistor
MOSFET:	Metal-Oxide-Semiconductor Field-Effect-Transistor
OSD:	Open/Short De-embedding
PA:	Power Amplifier
PCS:	Personal Communication Services
PCN:	Personal Communication Network
RF:	Radio Frequency
RL:	Return Loss
RMS:	Root-Mean-Square
SFDR:	Spurious-Free Dynamic Range
SG:	Signal-Ground
SNR:	Signal-to-Noise Ratio

SOLT:	Short-Open-Load-Through
SS:	Signal-Signal
VSWR:	Voltage Standing Wave Ratio

Chapter 1

Introduction

The world is undergoing a transformation to a next generation of communications systems. Wireless technology plays a key role in these new technologies and devices which include Personal Communication Services (PCS), mobile data, paging, wireless Local Area Networks (LANs) [1].

With the increasingly widespread use of portable wireless communications devices such as cellular phones, message pagers, and wireless modems, the demand for low-cost, low-power, small-size and highly integrated RF transceivers increases dramatically as well.

1.1 Architecture of an RF System

The basic structure of an RF transceiver system for personal wireless communications is shown in Figure 1.1. In the transmitter path, the voice signal is first digitized by an analog-to-digital converter (ADC), and then applied to the modulator, modulating an RF frequency carrier. Subsequently, the output data is amplified by a power amplifier (PA) so as to drive the antenna. The antenna radiates the signal into the air, and the transmission is completed.

In the receiver path, the basic operation is simply inverse. The signal received by the antenna is amplified by a low-noise amplifier (LNA), and then the spectrum is translated to a baseband frequency by a mixer. After being demodulated, the signal is then converted to analog form by a digital-to-analog converter (DAC), amplified, and applied to the speaker [2].

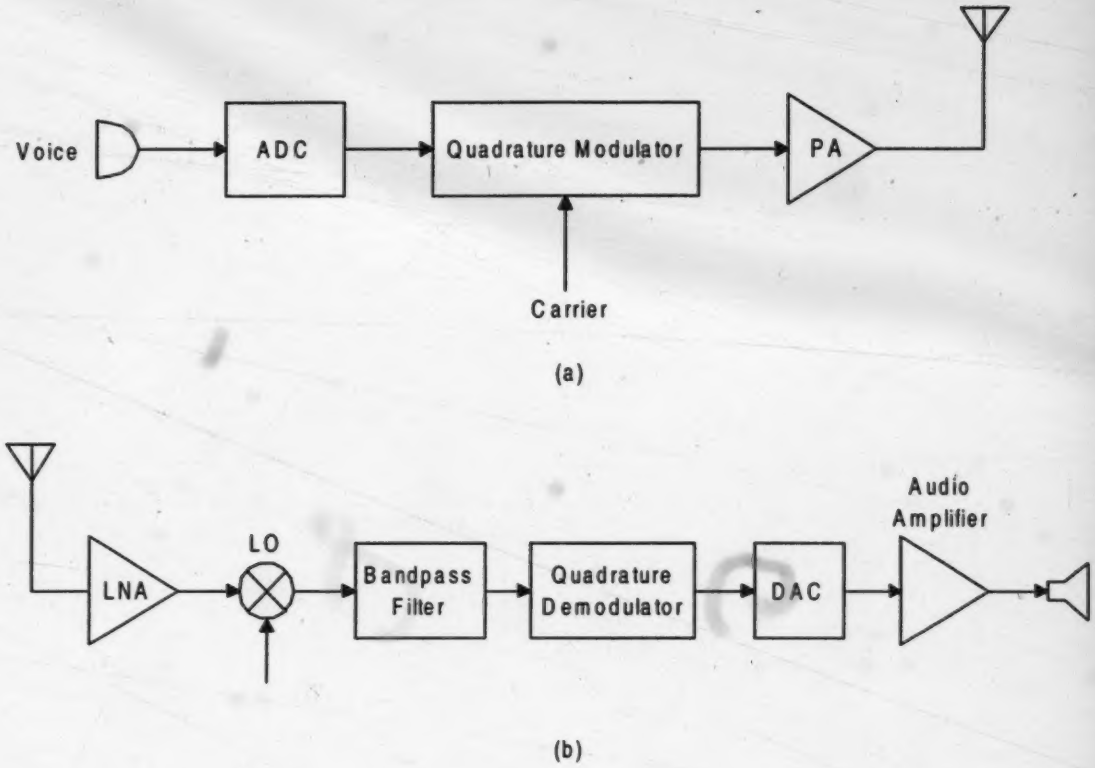


Figure 1.1: Block diagram of a common RF transceiver: (a) transmitter, (b) receiver.

1.2 Low Noise Amplifier Fundamentals

The LNA in the receiver path is the most basic and important cell for an RF transceiver. The main objective of the LNA circuit is to amplify the low power RF signals that appear at the antenna, in order for the rest of the receiver subsystem to be able to process them efficiently [4].

In this section, some common characteristics of LNAs are defined: noise figure, gain, bandwidth, linearity, input third-order intercept point (IIP3), input and output impedance, input and output return loss, and stability factor. Frequently used design equations are also summarized.

1.2.1 Noise Figure

Noise figure (NF) is commonly used to measure the noise performance of an amplifier. Noise figure is defined as ten times the logarithm (base 10) of the noise factor (F). The noise factor is defined as

$$F = \frac{\text{total output noise}}{\text{total output noise due to the source}}.$$

The noise figure of a two-port network can also be expressed as the degradation of the signal-to-noise ratio (SNR) between the input and output ports. That is,

$$NF = \frac{\text{input } SNR}{\text{output } SNR} = \frac{P_{si}/P_{ni}}{P_{so}/P_{no}} = \frac{P_{no}}{G_a P_{ni}} = 1 + \frac{P_{ne}}{G_a P_{ni}},$$

where P_{si} and P_{ni} represent input signal and noise powers respectively, P_{so} and P_{no} represent corresponding output signal and noise powers, G_a is the frequency-dependent available power gain of the network and P_{ne} is output noise power due to internal noise sources.

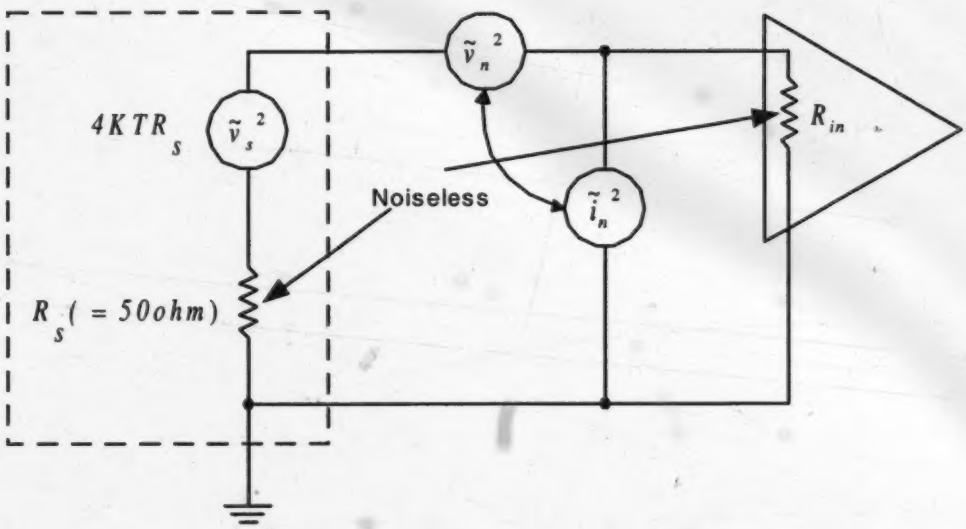


Figure 1.2: Noise figure calculation for an amplifier.

In the model of Figure 1.2, all of the noise sources appear as inputs to the noiseless amplifier, where $\tilde{v}_n^2$ and $\tilde{i}_n^2$ are the amplifier input-referred noise-voltage and noise-

current generators, respectively. Then, the expression for noise figure becomes [5]

$$(NF)_{dB} = 10\log\left[1 + \frac{V_n^2 + 2\rho V_n I_n R_s + I_n^2 R_s^2}{4kTR_s\Delta f}\right], \tag{1.1}$$

where k is Boltzmann's constant (about $1.38 \times 10^{-23} J/K$), T is the absolute temperature in Kelvin, R_s is the source resistance, Δf is the bandwidth in hertz over which the noise is measured, ρ is the correlation coefficient between $\tilde{v}_n^2$ and $\tilde{i}_n^2$, and V_n and I_n are the root-mean-square (RMS) noise voltage and current, respectively. The noise power is proportional to the square of V_n [5]. Note that the noise figure is strongly influenced by R_s, not dependent on the input impedance, i.e., matching of the amplifier. For high R_s, $\tilde{i}_n^2$ is the dominant equivalent noise source and conversely, for low R_s, $\tilde{v}_n^2$ is the dominant equivalent noise source [3].

From equation (1.1), NF can be found to be minimum when

$$R_{s,opt}^2 = \frac{\tilde{v}_n^2}{\tilde{i}_n^2},$$

assuming negligible correlation between $\tilde{v}_n^2$ and $\tilde{i}_n^2$ [6]. As the biasing of the LNA input transistor is varied, the noise figure varies. Thus there is an indirect relationship between NF and input matching.

Primary interest in the evaluation of noise performance of multistage amplifiers is the overall noise figure of the system. The cascaded noise figure for an n-stage system is given by [7]

$$NF_{tot} = NF_1 + \frac{NF_2 - 1}{G_{a1}} + \frac{NF_3 - 1}{G_{a1}G_{a2}} + \cdots + \frac{NF_n - 1}{G_{a1}G_{a2}\ldots G_{a(n-1)}}, \tag{1.2}$$

where NF_i is the linear noise figure of stage i and G_{ai} is the linear available power gain of the stage i which is obtained with a matched load [9].

It is obvious from equation (1.2) that the NF of the first stage in the n-stage system has the predominant effect on the overall NF, unless G_{a1} is small or NF_2 is large. Therefore, the noise produced in the first stage or two should be minimized by the choice of low-noise transistors and the selection of operating conditions. As we know, the LNA is the first active circuit in the receiver following the antenna; it dominates the noise performance of the receiver. Hence, it must exhibit as low noise as possible.

1.2.2 Gain

In addition to noise figure, power gain (G) is another most important consideration in the LNA system. The power gain of an amplifier is the ratio of the output power (P_o) to the input power (P_i). It can be written as [8]

$$G = \frac{P_o}{P_i}. \tag{1.3}$$

The decibel power gain is defined as

$$G_{dB} = 10 \log G.$$

Assuming that the input impedance (R_i) and load (R_L) are purely resistive, equation (1.3) becomes

$$G = \frac{P_o}{P_i} = \frac{V_o I_o}{V_i I_i} = A_v A_i = A_v \frac{v_o/R_L}{v_i/R_i} = (A_v)^2 \frac{R_i}{R_L}, \tag{1.4}$$

where uppercase symbols, such as V_o and I_i, stand for the root-mean-square (RMS) values of the voltages and currents; lowercase symbols v_o and v_i, stand for the instantaneous values; A_i and A_v represent current and voltage gain respectively.

For a cascade connection of the amplifiers, the overall voltage gain is given by

$$A_{v,tot} = A_{v1} A_{v2} ... A_{vn}. \tag{1.5}$$

Thus, the overall voltage gain of cascaded amplifier stages is the product of the voltage gains of the individual stages. Similarly, the overall current gain of a cascaded connection of amplifiers is the product of the current gains of the individual stages. Furthermore, the overall power gain is the product of the individual power gains, which can be described as

$$G_{tot} = G_1 G_2 ... G_n. \tag{1.6}$$

In order to minimize the noise contribution from the rest of the stages in the receiver, especially that from the mixer, the LNA must provide as high gain as possible with no distortion introduced and, at the same time, add as little noise as possible.

1.2.3 Bandwidth

The half-power (-3 dB) bandwidth is an important specification for amplifiers. It is the usable frequency range of an amplifier across which the circuit can deliver a given

voltage or power output. Usually, the voltage bandwidth is not the same as the power bandwidth [3].

The bandwidth is defined as

$$BW = f_H - f_L,$$

where f_L and f_H are the half-power (or -3 dB) frequencies, that is, the frequencies at which the voltage (or current) gain is approximately 3 dB lower than the midband gain whose magnitude is constant over a wide range of frequencies known as the midband region.

There is a tradeoff between bandwidth and gain. The bandwidth increases while reducing gain. Sometimes, gain-bandwidth product (GBP) is used as a figure of merit for the performance of an amplifier [21]. For an amplifier with a single pole, the gain-bandwidth product is constant.

In an n-stage amplifier, the gain-bandwidth product is defined as [3]

$$(GBP)_n = (Gain)^{1/n} \cdot (BW), \tag{1.7}$$

where $Gain$ is the total midband gain, and BW is the overall bandwidth.

For an amplifier that is a cascade of n identical stages each with a single-pole response given as

$$\frac{v_o}{v_i} = \frac{G}{1 + j(\frac{\omega}{B})},$$

the overall transfer function is

$$H(j\omega) = \frac{G^n}{(1 + j\frac{\omega}{B})^n}. \tag{1.8}$$

Here G and B are the gain and bandwidth of each individual stage, respectively. Therefore, the gain-bandwidth product of each stage is $G \cdot B$.

From equation(1.8), it is easy to find that the total midband gain is $Gain = G^n$, while the overall bandwidth is when $|H(j\omega)| = \frac{G^n}{\sqrt{2}}$. This gives

$$BW = B\sqrt{2^{\frac{1}{n}} - 1}. \tag{1.9}$$

Obviously, as the number of stages increases, the overall bandwidth decreases.

Using equation(1.7), the amplifier gain-bandwidth product is

$$(GBP)_n = G \cdot B\sqrt{2^{\frac{1}{n}} - 1}. \tag{1.10}$$

So, as n increases, the amplifier gain-bandwidth product is reduced.

1.2.4 Linearity

When the LNA receives a weak signal in the presence of a strong interfering one, the strong signal can overload the LNA, lead to distortion, and generate new frequencies which may swamp out the desired weak signal. Thus, the linearity of an LNA determines the maximum allowable input signal level. The most commonly used measures of linearity are 1-dB compression point (P_{1dB}), third-order intercept (IP3) and dynamic range (DR).

Writing the amplifier output-to-input transfer characteristic as a polynomial,

$$v_o = a_0 + a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + \cdots = \sum_{n=0}^{\infty} a_n v_i^n, \tag{1.11}$$

where v_i is the instantaneous input voltage, and the coefficients a_n are assumed to be frequency independent. The output DC offset level is represented by a_0, $a_1 v_i$ is the first-order (linear) term, i.e., the small signal gain, $a_2 v_i^2$ is the second-order (square-law) term, and so on.

For a sinusoidal input $v_i(t) = A\cos(\omega t)$,

$$v_o(t) = a_0 + a_1 A\cos(\omega t) + a_2 A^2 \cos^2(\omega t) + a_3 A^3 \cos^3(\omega t) + \cdots.$$

Performing the algebra using trigonometric properties [4],

$$\begin{aligned} v_o(t) \;=\; & (a_0 + \frac{1}{2}a_2 A^2 + \frac{3}{8}a_4 A^4 + \cdots) + (a_1 A + \frac{3}{4}a_3 A^3 + \cdots)\cos(\omega t) + \\ & (\frac{1}{2}a_2 A^2 + \frac{1}{2}a_4 A^4 + \cdots)\cos(2\omega t) + (\frac{1}{4}a_3 A^3 + \cdots)\cos(3\omega t) + \cdots \end{aligned} \tag{1.12}$$

From equation (1.12), we observe that if a pure sinusoidal signal is applied at the input of a nonlinear system, at the output of the system, apart from the fundamental frequency, its harmonic products will appear.

However, it is possible for signals at various frequencies to appear at the input of a telecommunications system. This causes the formation of intermodulation (IM) products at the output. Assume that an input signal is formed by two sinusoidal components:

$$v_i(t) = A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t). \tag{1.13}$$

Substituting equation (1.13) into $v_o(t) = a_1 v_i + a_2 v_i^2 + a_3 v_i^3$, and after collecting terms, at the fundamental frequencies we have:

$$\text{At } \omega_1 : \quad (a_1 + \frac{3}{2} a_3 A_2^2 + \frac{3}{4} a_3 A_1^2) A_1 \cos(\omega_1 t)$$

$$\text{At } \omega_2 : \quad (a_1 + \frac{3}{2} a_3 A_1^2 + \frac{3}{4} a_3 A_2^2) A_2 \cos(\omega_2 t)$$

Second-order terms:

$$\text{At } 2\omega_1 : \quad (\frac{1}{2} a_2 A_1^2) \cos(2\omega_1 t)$$

$$\text{At } 2\omega_2 : \quad (\frac{1}{2} a_2 A_2^2) \cos(2\omega_2 t)$$

$$\text{At } \omega_1 \pm \omega_2 : \quad (a_2 A_1 A_2)[\cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t] \tag{1.14}$$

Third-order terms:

$$\text{At } 3\omega_1 : \quad (\frac{1}{4} a_3 A_1^3) \cos(3\omega_1 t)$$

$$\text{At } 3\omega_2 : \quad (\frac{1}{4} a_3 A_2^3) \cos(3\omega_2 t)$$

$$\text{At } 2\omega_1 \pm \omega_2 : \quad \frac{3}{4}(a_3 A_1^2 A_2)[\cos(2\omega_1 + \omega_2)t + \cos(2\omega_1 - \omega_2)t] \tag{1.15}$$

$$\text{At } \omega_1 \pm 2\omega_2 : \quad \frac{3}{4}(a_3 A_1 A_2^2)[\cos(\omega_1 + 2\omega_2)t + \cos(\omega_1 - 2\omega_2)t] \tag{1.16}$$

The terms in equation (1.14) are called second-order intermodulation products. Similarly, the terms in equations (1.15) and (1.16) are third-order intermodulation products. The intermodulation distortion (IMD) coefficients IM_i are defined as the ratio of the amplitude of each intermodulation product to the amplitude of the fundamental component at the output of the system.

The output power level where the fundamental frequency component deviates from its ideally expected value by 1 dB, is called 1-dB compression point [4]. The 1-dB compression point is typically measured in dBm. IP3 is the point at which the amplitude of the third-order intermodulation distortion (IM_3) would become equal to that of the fundamental. The input power level at which this occurs is called the input third-order intercept point (IIP3). Dynamic range (DR) is generally defined as the ratio of maximum allowable input level to minimum detectable input level . This definition is quantified differently in different applications. If we base the definition of the upper end of the dynamic range on the intermodulation behavior and the lower

end on the sensitivity, such a definition is called the spurious-free dynamic range (SFDR) [3]. The SFDR represents the maximum relative level of interferers that an LNA can tolerate while producing an acceptable signal quality from a small input level.

The definitions of both 1-dB compression point and IIP3 are graphically shown in Figure 1.3.

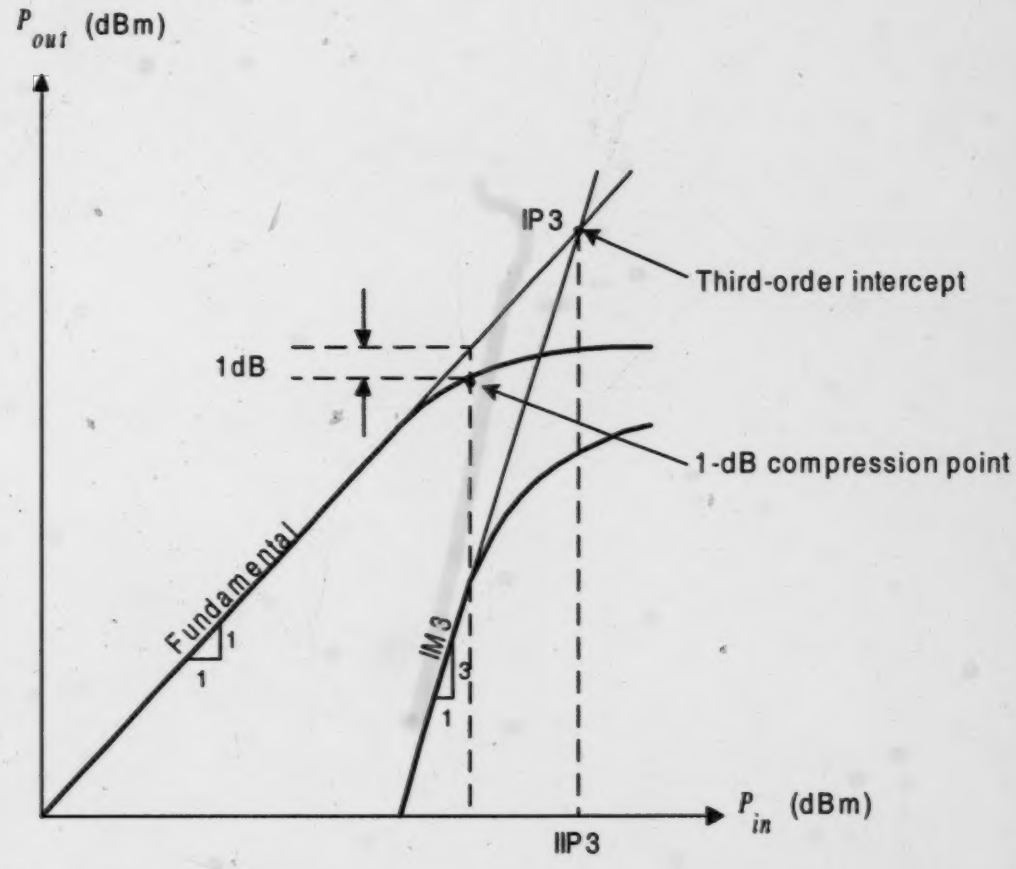


Figure 1.3: Graphical representation of 1-dB compression point, IIP3.

1.2.5 Input and Output Matching

For maximum power transfer and elimination of the standing waves, it is demanded to obtain matching between the antenna and the LNA, i.e., the LNA is designed to have a 50-Ω resistive input impedance in a frequency range as broad as possible. Notice that 50 Ω is a standard termination impedance level. Similarly, to drive the image-reject filter with minimum loss and ripple, the LNA output impedance must also be equal to 50 Ω.

A common measure of impedance matching is the voltage standing wave ratio (VSWR) which is defined as

$$VSWR = \frac{1 + |\Gamma|}{1 - |\Gamma|},$$

where Γ is the reflection coefficient:

$$\Gamma = \frac{Z - Z_0}{Z + Z_0}. \tag{1.17}$$

Here Z is input/output impedance of LNA, and Z_0 is source/load impedance which generally takes the value as 50 Ω. Return loss (RL) is defined (in dB) as

$$(RL)_{dB} = -20 \log |\Gamma|. \tag{1.18}$$

It is apparent that in the case of perfect matching, $VSWR = 1$ and $RL = 0$, or $(RL)_{dB} = \infty\ dB$.

1.3 CMOS Fundamentals

In this section, some MOS design equations are briefly summarized, and the comparison between CMOS and bipolar devices for their use in RF applications is described as well.

1.3.1 Basic Equations

When an NMOS transistor is in the saturation region, that is

$$V_{gs} - V_{ds} \leq V_t,$$

where V_{gs} represents gate-to-source voltage, V_{ds} represents drain-to-source voltage, and V_t represents threshold voltage; the drain current of the saturated NMOS transistor is given below:

$$I_D = \frac{\mu_n C_{ox}}{2} \frac{W}{L} (V_{gs} - V_t)^2, \tag{1.19}$$

ignoring the effect of V_{ds}. Here μ_n is mobility, C_{ox} is gate-oxide capacitance per unit area, W is the width of the channel, and L is the length of the channel.

The transconductance is easily found from differentiating the expression for drain current

$$g_m = \frac{\partial I_D}{\partial V_{gs}} = \mu_n C_{ox} \frac{W}{L} (V_{gs} - V_t) = \sqrt{2 \mu_n C_{ox} \frac{W}{L} I_D}. \tag{1.20}$$

The unity gain frequency, ω_T, of a MOSFET is

$$\omega_T = \frac{g_m}{C_{gs}} = \frac{\mu_n C_{ox} (W/L) (V_{gs} - V_t)}{\frac{2}{3} W L C_{ox}} = \frac{3}{2} \frac{\mu_n (V_{gs} - V_t)}{L^2}, \tag{1.21}$$

where C_{gs} is the capacitance between the gate and the source.

Notice that equations (1.19), (1.20) and (1.21) hold only in the long-channel regime. If the effective channel length (L) which is modulated by V_{ds} decreases, the expressions of I_D, g_m and ω_T are modified as follows [9]:

$$I_D = \frac{\mu_n C_{ox}}{2} W (V_{gs} - V_t) E_{sat}, \tag{1.22}$$

$$g_m = \frac{\partial I_D}{\partial V_{gs}} = \frac{\mu_n C_{ox}}{2} W E_{sat}, \tag{1.23}$$

$$\omega_T = \frac{g_m}{C_{gs}} = \frac{\frac{1}{2} \mu_n C_{ox} W E_{sat}}{\frac{2}{3} W L C_{ox}} = \frac{3}{4} \frac{\mu_n E_{sat}}{L}, \tag{1.24}$$

where E_{sat} is the field strength at which the carrier velocity has dropped to one half the value extrapolated from low-field mobility. A typical value for E_{sat} is about 4×10^6 V/m. From equation (1.24), we see that with the reduction of the minimum effective channel length, it is easy for unity gain frequency (f_T) to reach the range of 30-50 GHz, which is similar to that offered by many high-performance bipolar processes [9]. This is one reason that why the CMOS devices are increasingly chosen in RF applications.

1.3.2 Noise Sources in CMOS and Bipolar Devices

For MOS devices, the flicker noise prevails at low frequencies, whereas the thermal noise is dominant at high frequencies. Note that these two noise components are uncorrelated. Therefore, for RF operation, the dominant noise source is the thermal noise generated in the transistor channel. This noise is white with a power spectral density given by [10]

$$\frac{\bar{i_t^2}}{\Delta f} = 4kT\gamma g_{d0}, \tag{1.25}$$

where $\bar{i_t^2}$ is the shunt noise current source, g_{d0} is the zero-bias drain conductance of the device, and γ is a bias-dependent coefficient taking values as 2/3 for long-channel devices, and as high as 2 or 3 for short-channel devices.

The main noise sources in the bipolar transistor are thermal noise in the base spreading resistance, shot noise in the base and collector bias current [6], [11]. Obviously, CMOS devices have less white noise sources than bipolar devices at radio frequencies; the noise intensity is important in noise analysis.

From the above comparisons, CMOS devices are deemed to be suitable for RF applications. CMOS devices can offer high linearity, low noise and fast speed of operation at RF when they are carefully designed and laid out.

1.4 Thesis Organization

The thesis is organized as follows:

Chapter 2 will present the background of the LNA circuit design. Several LNA architectures and wideband amplification methods will be reviewed. Basic wideband LNA concepts and design equations will also be addressed.

Chapter 3 will discuss the analytical models for on-chip inductors and transformers. Several different configurations of transformers will be illustrated. The design guidelines of a spiral inductor will be given as well. The simulation and measurement results will be summarized.

Chapter 4 will discuss the overall design requirements for the RF wideband and low noise amplifier in this project. The CMOS process will be compared with other

technologies. The multi-stage design topology chosen in this design will be discussed as well.

Chapter 5 will discuss the implementation of the ideas presented in Chapter 4, as well as other circuit design issues. The layouts of the design will be given in detail and the simulations of the designed wideband LNA will be presented at the end of this chapter.

Chapter 6 will discuss the testing equipment, testing procedures and measurement results that were obtained from the design.

Finally, conclusions from this work, improvement and future work will be presented in Chapter 7.

Chapter 2

Wideband Amplification and Matching

2.1 Introduction

An LNA in the RF front-end of a wireless receiver must meet several exacting specifications. First is a low noise contribution, that is, NF $<$ 2.0 dB. The input noise of an LNA must be sufficiently low in order to detect weak input signals. Second, a high LNA gain is desirable to overcome the noise of later circuits (i.e., mixer), but not so much to cause mixer overload. Third, the input impedance of an LNA must be matched to the source impedance to reduce the effect of reflections. Usually, an LNA input impedance is designed to be 50 Ω.

Higher frequency bands are being sought to avoid congestion. Therefore, wireless systems are requiring higher frequency and wider bandwidth amplification [3]. For an LNA, wideband behavior will also be performed. For example, the operating frequency bands of mobile and wireless communications systems such as Advanced Mobile Phone Service (AMPS) and Pan European Group Special Mobile (GSM) are 900 MHz, and 1.9 GHz for the Personal Communication Network (PCN) and Digital European Cordless Telephone (DECT) [31].

2.2 Wideband Amplification Methods

Wideband amplifiers have constant gain from low frequency (sometimes down to DC) to very high frequency [8]. Since operating frequencies of many commercial wireless systems are in the low gigahertz range, amplifiers must provide stable gain at these

frequencies [3].

Negative feedback is a common technique used in wideband amplification to exchange gain for bandwidth and to optimize the high-frequency performance. The main benefits with negative feedback include: stabilization of gain, reduction of nonlinear distortion, reduction of certain types of noise, control of the input and output impedance and extension of bandwidth. However, these benefits of negative feedback are accompanied by two disadvantages. First, the gain of an amplifier is reduced. It can be overcome by adding extra amplifier stages of amplification. The second potential problem is the tendency for oscillation to occur in the circuit, which requires careful attention of the designer.

2.2.1 Local Feedback Configuration

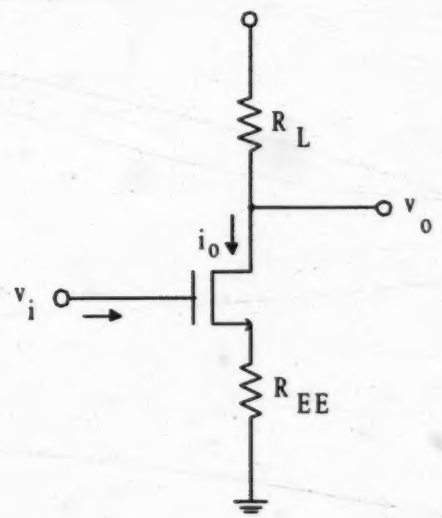


Figure 2.1: Transconductance configuration.

Local feedback is a direct carryover from discrete design techniques where each individual feedback stage can be independently biased and AC coupled to each other [32]. First we consider common-source configurations with one feedback element. Figure 2.1 shows the configuration featuring series feedback at input and output. Such a configuration has high input and output impedance. This circuit is suitable

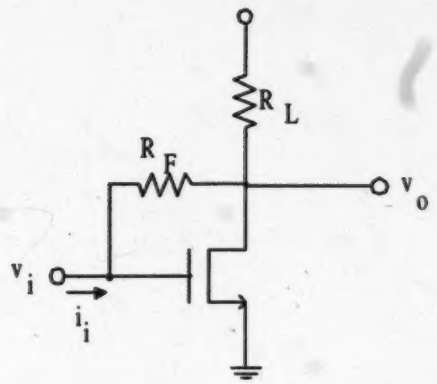


Figure 2.2: Transimpedance configuration.

to be driven by a voltage source, and the output behaves as a current source that is controlled by the input voltage. Therefore, this configuration works as a transconductance amplifier: $i_o = G_t v_i$. Here G_t is the transconductance, which is given by

$$G_t \approx \frac{1}{R_{EE}}. \tag{2.1}$$

Figure 2.2 shows the local feedback configuration incorporating parallel feedback. This configuration features low input and output impedance. Hence, it is particularly suited to be current-driven, whereas the output has the characteristics of a voltage source. Therefore, this configuration works as a transimpedance amplifier: $v_o = R_t i_i$. Here R_t is the transimpedance, which is given by

$$R_t \approx -R_F. \tag{2.2}$$

Both local feedback configurations are very useful building blocks in wideband amplifier design, especially when they are cascaded. Figure 2.3 shows the transconductance-transimpedance cascade. This combination features high input impedance and low output impedance which is characterized as a voltage amplifier. From equations (2.1) and (2.2), the voltage gain can be obtained

$$A_v \approx \frac{R_F}{R_{EE}}. \tag{2.3}$$

Figure 2.4 shows the reverse combination: transimpedance-transconductance cascade. This configuration has low input impedance and high output impedance, which

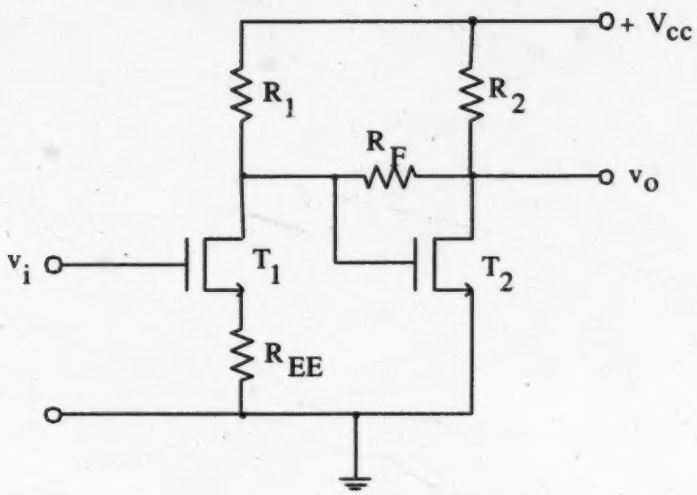


Figure 2.3: Voltage amplifier.

is typical for a current amplifier. Similarly, from equations (2.1) and (2.2), the current gain can be expressed as

$$A_i \approx \frac{R_F}{R_{EE}}. \tag{2.4}$$

2.2.2 Overall Feedback Configuration

An alternate approach to the design of wideband amplifiers is the use of overall feedback around a number of gain stages, rather than applying local feedback around each individual stage. As compared with the local feedback cascades described above, the overall feedback offers a higher degree of bias stability and desensitivity to individual gain tolerances [32].

Figures 2.5 and 2.6 show the basic structures of two-stage global feedback configurations often used in wideband amplifier design. The circuit of Figure 2.5 features high input impedance and low output impedance, which characterizes the configuration as a typical voltage amplifier. On the contrast, the circuit of Figure 2.6 has low input impedance and high output impedance, which makes it a good current amplifier.

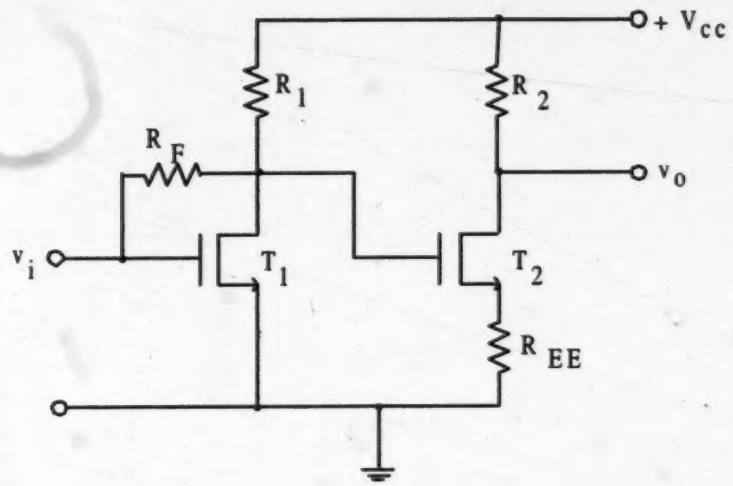


Figure 2.4: Current amplifier.

2.2.3 Compound Transistors

Some improvement of bandwidth can be also achieved using compound transistors. The term "compound" means a combination of two or three transistors, acting together as a single active component with improved behavior. The most commonly used compound transistor configurations in wideband amplifier design are common-drain-common-source (CD-CS) and common-source-common-gate (CS-CG) connections.

The common-drain-common-source configuration, referred to as a Darlington configuration [53] is shown in Figure 2.7. The resistor, R, is necessary to bias the transistor T_1. Without R, the source current of T_1 will be equal to the gate current of T_2 which is very small. We know that low transistor current results in low transistor f_T. In other words, T_1 is biased to operate with a relatively high drain current, comparable to that of T_2, in order to optimize its f_T [32]. However, the resistor R does more than biasing the transistor, it shunts some of the RF signal to ground, and affects the AC response of the circuit.

The Darlington configuration features high input impedance, high DC current gain, and the ability to amplify beyond f_T. However, it exhibits poor phase charac-

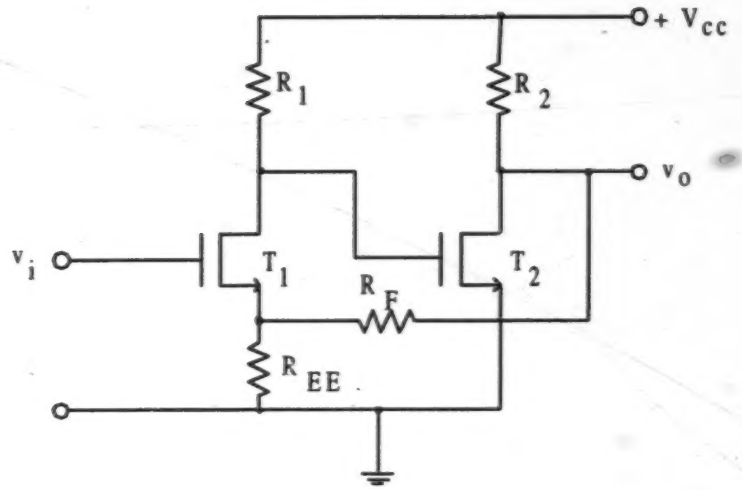


Figure 2.5: Two-stage voltage amplifier with global feedback.

teristics, which may cause instability. With decreasing the value of R, the gain and phase-shift both drop, and therefore increases the stability.

The common-source-co mon-gate configuration is commonly referred to as the cascode configuration. A schematic of a CMOS version , shown in Figure 2.8, consists of a common-source stage driving a common-gate stage. The cascode derives its advantage at high-frequency from the fact that the drain load for the common-source stage is the very low input impedance of the common-gate stage [6]. If T_1 and T_2 have equal drain bias currents, the transconductance of T_1 is g_m, and, since its load resistance is $1/g_m$, the voltage gain of T_1 is unity; the voltage gain of the combination is $g_m R_L$. Thus the influence of the Miller effect on T_1 is minimal even for large values of R_L. As a result, the cascode connection has better high-frequency performance than the single common-source stage, especially for large R_L.

The cascode configuration exhibits a higher degree of isolation (i.e., a small amount of reverse signal transmission) between its input and output terminals than a simple common-source transistor [32]. Another useful characteristic of the cascode is its high output impedance, which can be used to advantage in current-source design.

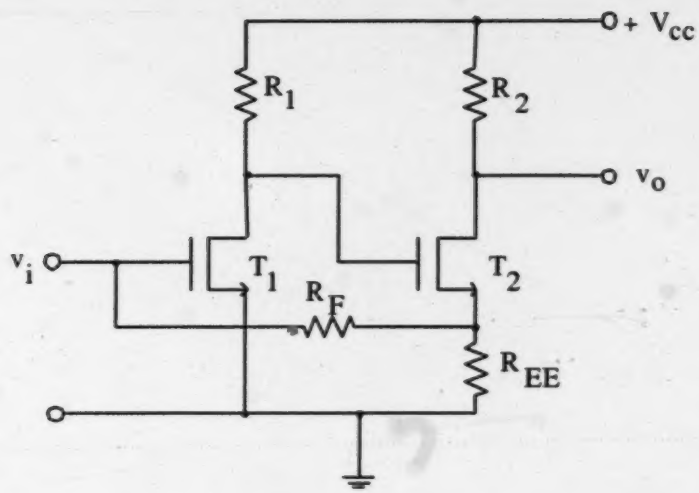


Figure 2.6: Two-stage current amplifier with global feedback.

2.2.4 Inductive Peaking

The relationship between bandwidth and gain has been discussed in Chapter 1. For all the basic amplifier stage configurations which can be characterized by one dominant pole, the gain-bandwidth product (GBP) is approximately the same, which is given by $g_m/C_{tot} \approx \omega_T$, affected mainly by the active device. Given a value of GBP, it is obvious that bandwidth can be traded for gain. Some amplifiers are performance-limited mainly by speed. The speed of an amplifier can be expressed by its response to a unit step input [33]. For a single-pole amplifier with pole $p = 1/\tau$, where τ denotes the time constant, one important quantity that characterized speed is t_r, which is the 10 to 90 % risetime in response to a unit step input. Here t_r is related to τ as

$$t_r = \tau \ln(9) \simeq 2.2\tau = \frac{2.2}{p}. \tag{2.5}$$

Considering the basic CS-stage used as a voltage amplifier shown in Figure 2.9, we can easily find the relationship between bandwidth and risetime.

As shown in Figure 2.9, the gain of this stage can be expressed as

$$A_{v1}(s) = -g_m R_L \cdot \frac{1}{1 + s\tau}, \tag{2.6}$$

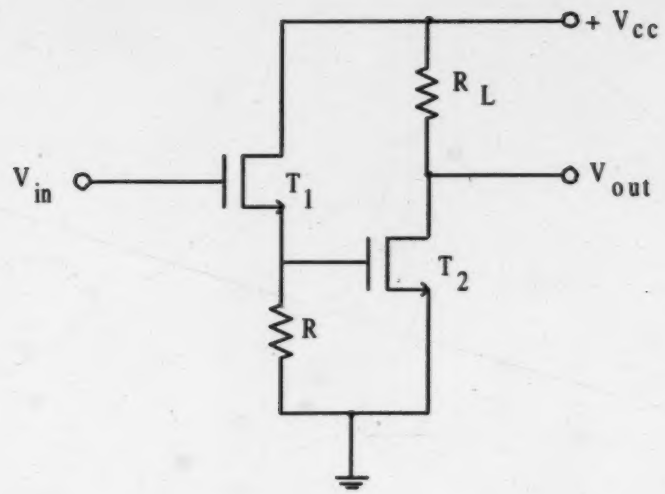


Figure 2.7: Common-drain-common-source configuration.

where R_L is resistive load and $\tau = R_L C_L$ with $C_L = C_{gs} + C_{ds}$, assuming that the influence of C_{gd} can be neglected. The -3 dB bandwidth in terms of angular frequency (in radians per second) is , to a first approximation, given by

$$\omega_{-3dB} = 1/\tau. \tag{2.7}$$

Substituting equation (2.5) into equation (2.7), we get

$$\omega_{-3dB} \cdot t_r \simeq 2.2. \tag{2.8}$$

Hence, the bandwidth-risetime product is about 2.2 for a single-pole system. As the risetime, t_r, decreases, the bandwidth increases. In other words, increasing the response speed can improve the bandwidth. Methods for achieving this are among the oldest techniques for wideband amplification.

Figure 2.10 shows a simple configuration called inductive shunt-peaking. This amplifier is a standard common-source configuration, with the addition of the inductance. Assuming that the transistor is ideal, the only elements that control the bandwidth are R_L, L, and C_L [9]. For a purely resistively loaded common-source amplifier, the gain is proportional to $g_m R_L$. When a capacitive load is added, the

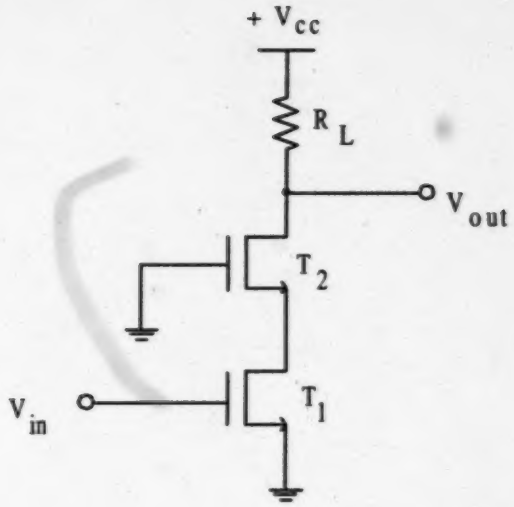


Figure 2.8: Cascode configuration.

gain gradually falls off as frequency increases because the capacitor's impedance diminishes. The addition of L in series with R_L is in parallel with C_L and creates a parallel resonance. For a step of input current, most of it charges C_L at first because the inductor delays current flow through the branch containing the resistor. Consequently, C_L charges faster and response speed increases. To the extent that a faster risetime or higher response speed implies a wider bandwidth, an appropriate choice of the inductor therefore increases the bandwidth.

The transfer function (v_{out}/i_{in}) of the shunt-peaking circuit of Figure 2.10 is just the impedance of the RLC network, which is written as

$$\begin{aligned} \frac{v_{out}}{i_{in}} = Z(s) &= (sL + R_L) \| \frac{1}{sC_L} \\ &= \frac{R_L + sL}{1 + sR_LC_L + s^2LC_L}. \end{aligned} \tag{2.9}$$

This transfer function includes two poles and one zero. Since the gain of the amplifier is the product of g_m and the magnitude of $Z(s)$, the latter can be computed as a

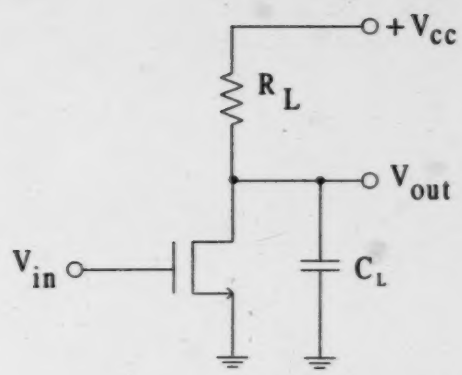


Figure 2.9: Basic CS-voltage amplifier configuration.

function of frequency as follows:

$$| Z(j\omega) | = R_L \sqrt{\frac{(\omega L/R_L)^2 + 1}{(1 - \omega^2 LC_L)^2 + (\omega R_L C_L)^2}}. \tag{2.10}$$

Notice that, compared with the simple RC case, there is a term in the numerator(from the zero) that increases with increasing frequency. Furthermore, the term of $1 - \omega^2 LC_L$ in the denominator contributes to an increase in $| Z |$ as the frequency increases. Obviously, adding L makes the high-frequency gain increase, whereas the low-frequency gain still does not change.

Figure 2.11 shows an inductive series-peaking circuit. The transfer function (v_{out}/i_{in}) for this amplifier is

$$\frac{v_{out}}{i_{in}} = \frac{R_L}{s^2 LC_L + sR_L C_L + 1}, \tag{2.11}$$

which has two complex conjugate poles, but no finite zeros. Hence, for control of the bandwidth, there is only one way, namely, damping of these poles. Usually, R_L is chosen to set the gain, C_L is parasitic and L is the design parameter. Compared to the shunt-peaking amplifier, the series-peaking amplifier is slower and has a narrower bandwidth. This kind of circuit is not used.

Two possible combinations of shunt and series peaking are shown in Figure 2.12. The operation of Figure 2.12(b) can be described as follows [9]. As in the step response of an ordinary shunt-peaking amplifier, the flow of current into the load resistor

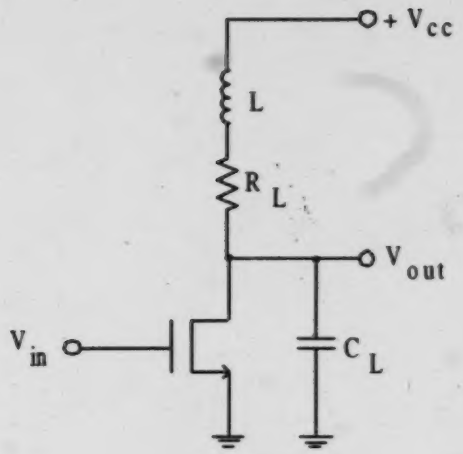


Figure 2.10: Shunt-peaking amplifier.

continues to be deferred by the action of L_1. This helps speed up the charging of the load capacitance. In addition, the transistor initially has to drive only its own output capacitance for some time, because L_3 delays the diversion of current into the rest of the network. Hence, risetime at the drain improves, which implies an improved bandwidth. Some time after the drain voltage has risen dramatically, the voltage across the load capacitance begins to rise as current finally starts to flow through L_2. Therefore, such a network charges the capacitances serially in time, rather than in parallel. The tradeoff is an increased delay in exchange for the improved bandwidth.

2.3 Common LNA Architectures

Design and fabrication of different LNA architectures in the 900 MHz-2 GHz frequency range using CMOS technology have been investigated by many researchers. Assuming that only the requirement of input impedance matching is focused on, four common CMOS LNA input stages will be introduced that are illustrated in Figure 2.13.

2.3.1 Common-Gate Architecture

When a noise figure of 3 dB is acceptable, it is simpler to control the input impedance with a common-gate input stage [27]. Common-gate circuit provides a low input resis-

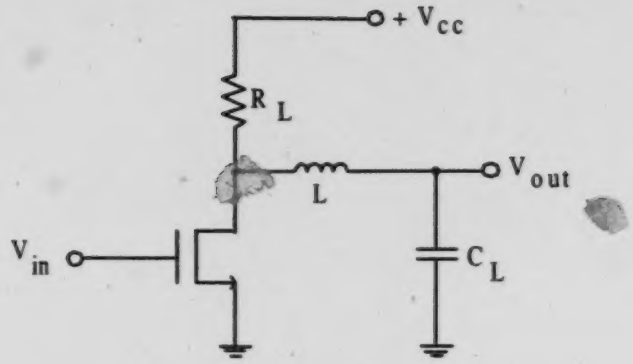


Figure 2.11: Series-peaking amplifier.

tance and a voltage gain which has a positive sign [21]. For this configuration shown in Figure 2.13(a), the input matching condition requires that the input resistance at the FET source, $1/g_m$, equals to R_s which is 50 Ω. Under such a matching condition, the noise factor (F) can be expressed as [10]

$$F = 1 + \frac{\gamma}{\alpha}, \tag{2.12}$$

where

$$\alpha = \frac{g_m}{g_{d0}}.$$

In the above equation, γ is the coefficient of channel thermal noise, g_m is the device transconductance, and g_{d0} is the zero-bias drain conductance. For long-channel devices, $\gamma = 2/3$ and $\alpha = 1$. Substituting these values into equation (2.12), we have $F = 5/3$ and then the noise figure $NF = 2.2$ dB. For short-channel devices γ can be much greater than one, and α can be much less than one. Accordingly, the minimum theoretically achievable noise figure tends to be around 3 dB or greater in practice.

2.3.2 Resistive Termination Architecture

This technique can provide a reasonably stable broadband input impedance. To achieve a good impedance match, a 50 Ω resistor, R_m, is simply put across the input terminals of a common-source amplifier; that is shown in Figure 2.13(b).

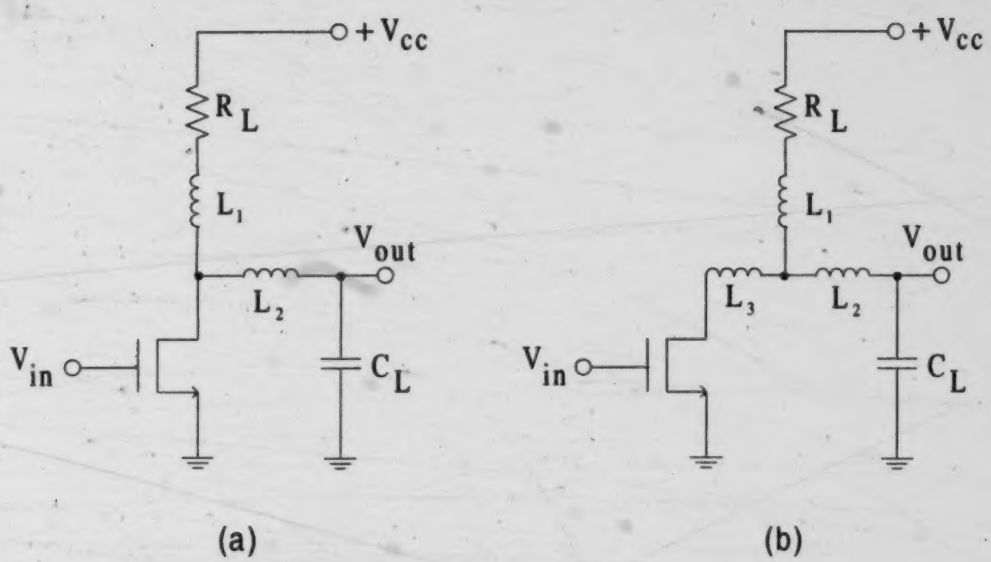


Figure 2.12: Shunt-series-peaking amplifiers.

The lower bound on the noise factor of this circuit can be expressed as [9]

$$F \geq 2 + \frac{4\gamma}{\alpha} \cdot \frac{1}{g_m R}, \tag{2.13}$$

where $R_m = R_s = R$. When g_m is high, the minimum noise figure can be found as

$$NF = 10 \log (2) = 3dB.$$

Apparently, the terminating resistor R_m adds thermal noise of its own and attenuates the signal ahead of the transistor as well. With the combination of these two effects, the noise figure is sharply degraded. Therefore, the large noise penalty makes this approach unattractive.

2.3.3 Shunt-Series Feedback Architecture

Figure 2.13(c) illustrates another configuration, which uses resistive shunt and series feedback to set the input and output impedances of the LNA. Formally, the input resistance R_{in} is given by [9]

$$R_{in} = \frac{R_F}{1 - A_V} \approx \frac{R_F}{1 + R_L/R_1}, \tag{2.14}$$

where A_V is the voltage gain from gate to drain. Similarly, the output resistance R_{out} can be expressed as

$$R_{out} = \frac{R_F + R_s}{1 + R_s/R_1} \approx \frac{R_F}{1 + R_s/R_1}. \tag{2.15}$$

If the source and load resistances are equal, that is $R_s = R_L = R$, then we may combine equation (2.14) and equation(2.15) into

$$R_{in} \approx R_{out} \approx \frac{R_F}{1 - A_V} \approx \frac{R_F}{1 + R/R_1}. \tag{2.16}$$

It is evident that the resistive shunt-series feedback continues to generate thermal noise of its own. Furthermore, amplifiers using feedback often have extraordinarily high power dissipation. Nonetheless, the advantage of the broadband capability of this circuit makes the shunt-series feedback amplifier to be found in many LNA applications, even though its noise figure is not the minimum possible.

2.3.4 Inductive Degeneration Architecture

For a common-source configuration shown in Figure 2.13(d), the real part of input impedance required for 50 Ω matching is generated by the inductive degeneration [26], [29]. The matching network consists of a series feedback inductor, L_s, in the FET source, another inductor, L_g, in series with the gate, and the gate-to-source capacitance, C_{gs}. This topology is preferred over resistor feedback because the matching network introduces no noise of its own. Practically, losses in inductors L_g and L_s will tend to degrade noise figure. However, with sufficiently good inductor, a noise figure below 3 dB may be obtained with this technique [27].

To simplify the analysis, consider a device model that includes only a transconductance, g_m, and a gate-to-source capacitance, C_{gs}. A small signal equivalent circuit of the inductively degenerated common-source amplifier shown in Figure 2.13(d) is illustrated in Figure 2.14.

It is not hard to establish the following form:

$$V_{in}(s) = I_{in}(sL_g) + I_{in}(\frac{1}{sC_{gs}}) + (I_{in} + g_m I_{in}(\frac{1}{sC_{gs}}))(sL_s). \tag{2.17}$$

Hence, the input impedance of the inductively degenerated common-source amplifier is

$$Z_{in}(s) = \frac{V_{in}(s)}{I_{in}(s)} = g_m(\frac{sL_s}{sC_{gs}}) + \frac{1}{sC_{gs}} + (sL_g + sL_s)$$
$$= g_m\frac{L_s}{C_{gs}} + \frac{1}{sC_{gs}} + s(L_g + L_s). \quad (2.18)$$

The first term in equation(2.18) is a frequency independent resistive term whose value can be set to R_s, 50 Ω, therefore the input impedance matching is achieved. The matching conditions can be shown [26] to occur when

$$\omega_0^2 C_{gs}(L_g + L_s) \approx 1, \quad (2.19)$$

$$L_s = R_s\frac{C_{gs}}{g_m}, \quad (2.20)$$

where ω_0 is the midband resonant frequency of operation. The inductor L_g is used to control the resonant frequency while L_s is constrained by one of the impedance matching criteria. Since the input impedance is purely resistive only at resonance, this approach can provide a narrow-band impedance matching only.

The effective transconductance of such an amplifier can be found by considering Figure 2.15.

$$I_{in} = \frac{V_s}{Z_{in'}},$$

$$I_{out} = I_{in} \cdot \frac{1}{sC_{gs}} \cdot g_m,$$

$$G_m = \frac{I_{out}}{V_s} = \frac{1}{sC_{gs}} \cdot \frac{g_m}{Z_{in'}}$$
$$= \frac{1}{sC_{gs}} \cdot \frac{g_m}{s(L_s + L_g) + \frac{1}{sC_{gs}} + \frac{g_m}{C_{gs}}L_s + R_s}$$
$$= \frac{g_m}{sC_{gs}(\frac{g_m}{C_{gs}}L_s + R_s)}. \quad (2.21)$$

Remember that $\omega_T = \frac{g_m}{C_{gs}}$. Under matching conditions, equation (2.21) can then be rewritten into

$$G_m = \frac{\omega_T}{2\omega_0 R_s}, \quad (2.22)$$

which is independent of the transistor g_m. The quality factor Q of the input circuit at resonance can be expressed as

$$\begin{aligned} Q = \frac{\omega_0 L}{R} &= \omega_0 \cdot \frac{(L_g + L_s)}{R_s + \frac{g_m L_s}{C_{gs}}} \\ &= \omega_0 \cdot \frac{(L_g + L_s)}{2R_s} \\ &= \frac{1}{2g_m \omega_0 L_s}, \end{aligned} \tag{2.23}$$

and the noise figure is given by

$$NF = 1 + \gamma \frac{1}{Q^2 \cdot g_m R_s}. \tag{2.24}$$

From equation (2.24), we see that choosing low g_m can improve noise figure. As g_m is reduced, Q which is related to g_m by equation (2.23) increases. A simple formula appeared in equation (1.20) $g_m = \sqrt{2\mu_n C_{ox} \frac{W}{L} I}$ is commonly used to estimate the current consumption. Since g_m is proportional to $\sqrt{I}$, the bias current I is inversely proportional to Q^2. As a result, a high Q matching network allows the current consumption of the LNA input stage to be low [30]. This is important for a CMOS LNA design. In summary, common-source LNAs with inductive source degeneration provide the lowest noise figure with low power consumption and a small device size.

2.4 Conclusion

The above discussion is of crucial importance in the study of LNAs. The different types of LNAs and what methods might be used to extend the bandwidth can be very important in the design process, as meeting the needed specifications is a critical concern. It should be emphasized that many bandwidth amplification techniques have not been used at the single-chip (CMOS) level, even though some research is going on for such a purpose.

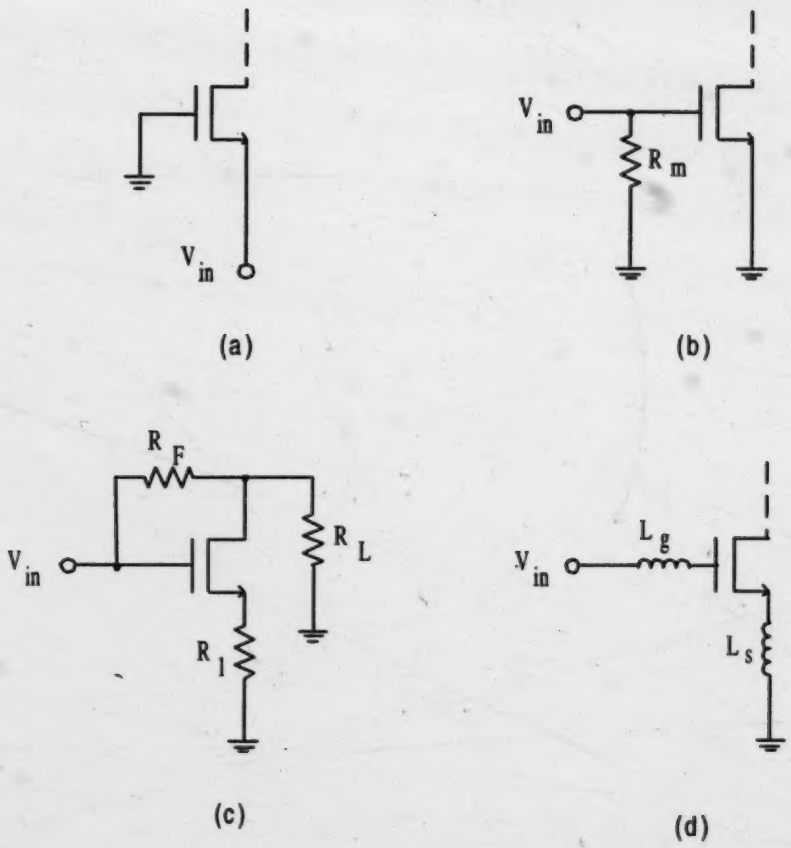


Figure 2.13: Common LNA architectures: (a) common-gate stage, (b) resistive termination, (c) shunt-series feedback, (d) inductive degeneration.

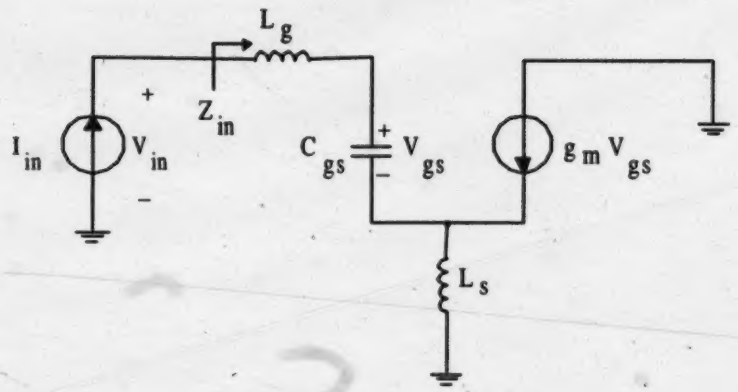


Figure 2.14: The equivalent circuit of the inductively degenerated common-source amplifier.

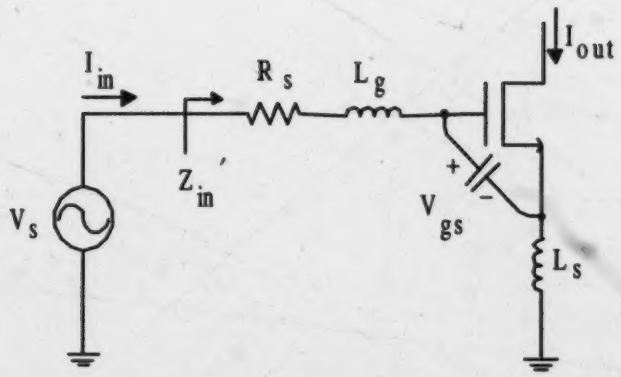


Figure 2.15: Schematic diagram used for effective transconductance calculations.

Chapter 3

Inductor and Transformer Design

3.1 Introduction

Spiral inductors and transformers are essential in silicon-based RF ICs. They have been widely applied in narrow-band impedance matching, tuned loads, low noise degeneration and feedback, linear filters, fully differential circuits and low voltage/low power design.

To design a good quality wideband LNA, which means with low noise, high frequency of operation, low distortion and low power dissipation, RF inductors and transformers are fundamental and indispensable components. However, it has been a challenge to integrate high-inductance inductors on silicon substrate for RF circuits operating at 1-2 GHz [13]. Two major factors which make it difficult to build large inductors with high self-resonant frequency on a silicon substrate are high substrate capacitance and high substrate loss. The high parasitic capacitance to the substrate lowers the self-resonant frequency, and the resistive loss in the substrate worsens the inductor Q-factor. Recent efforts have focused on increasing the substrate resistivity and dielectric thickness to increase inductor self-resonant frequency, and making use of multi-layer metals to reduce the series resistance in the metal windings and obtain higher Q [14].

This chapter describes a method for the analysis, modeling and optimization of spiral inductors and transformers. With the application of a custom CAD simulator ASITIC (Analysis and Simulation of Inductors and Transformers in Integrated Circuits), inductors and transformers are designed and laid out. The technology used is 0.35μm CMOS process with three metal layers of 0.07 ohm/sq for M1 and M2, and

0.04 ohm/sq for M3. The substrate is made of p-epi above a bulk layer.

3.2 Inductor Requirements

A practical inductor is specified by its inductance, quality factor (Q), and self-resonant frequency. With the increase of frequency, the impedance of a real inductor increases until self-resonant frequency is reached, where the parasitic capacitance resonates with the inductor. Beyond the resonant frequency, the inductor becomes capacitive and the impedance decreases with frequency.

For a good inductor, general requirements should be met which include large inductance, small series resistance, low substrate losses, small area, and high self-resonant frequency. A large inductance reduces the current consumption. A smaller series resistance with lower substrate losses improves the Q. A small area is desired to reduce the cost and a high self-resonant frequency enhances the operating range of the LNA.

3.3 Inductor and Transformer Modeling

3.3.1 Equivalent Circuits of Spiral Inductors

A typical spiral inductor is shown in Figure 3.1 with four layout parameters: the length of the outer segments A, metal width w, metal spacing s and a number of turns N. This spiral inductor consists of a number of series-connected metal segments. An equivalent frequency independent pi-circuit for the square-spiral inductor is shown in Figure 3.2 [16]. In this circuit, L models the self and mutual inductance in the segments, R is the accumulated sheet resistance, C_{s1} and C_{s2} model the parasitic capacitances from the metal layer to the substrate, and R_{s1} and R_{s2} represent the resistances of the conductive Si substrate.

A modification to Figure 3.2 is shown in Figure 3.3 which models second-order effects that increase the effective usable bandwidth of the model. For instance, C_{BR} has been added to model the interwinding capacitance and to make the series inductance and resistance frequency dependent; R_C is a substrate coupling resistor which models the substrate coupling impedance.

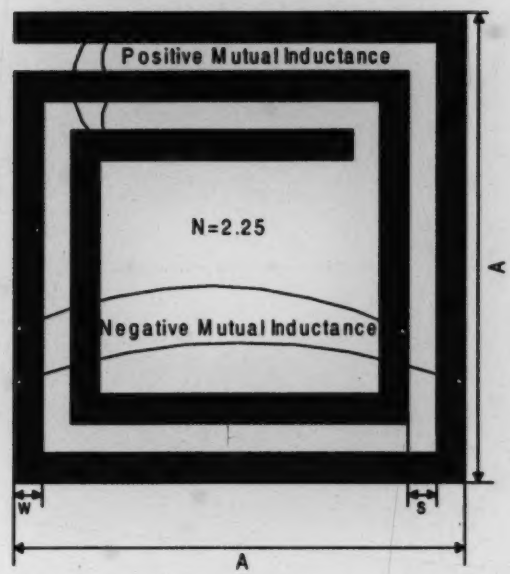


Figure 3.1: A spiral inductor layout.

3.3.2 Equivalent Circuit of Coupled Inductors

Similar to inductors, transformers are important elements in silicon RF designs to perform impedance conversion, impedance matching, bandwidth enhancement, signal coupling, phase splitting, etc. The desired characteristics for a transformer are application-dependent. Transformers can be configured as three-terminal or four-terminal devices. They may be used for narrow- or broad-band applications. For example, in single-ended-to-differential conversion, the transformer is used as a narrowband four-terminal device [17]. In this case, a high mutual coupling coefficient and high self-inductance are desired along with low series resistance. On the other hand, for bandwidth extension applications, the transformer is used as a broadband three-terminal device. In this case, a small mutual coupling coefficient and high series inductance is acceptable while all capacitances need to be minimized [18].

Two coupled inductors can be considered as a transformer. Assume that two spirals are placed on the silicon substrate as shown in Figure 3.4, with one side of each inductor being grounded. These inductors behave like a weakly coupled transformer and may be modeled as shown in Figure 3.5. It is a combination of inductor pi-

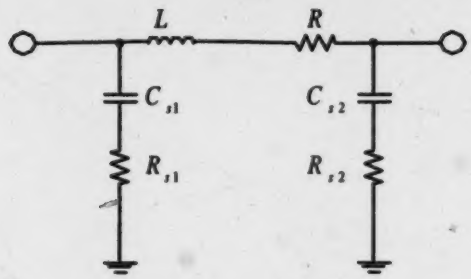


Figure 3.2: Traditional spiral inductor model.

model for each of the two segments plus coupling components between them. In this broadband model, R_C and $k_{1,2}$ represent the substrate coupling impedance and the mutual coupling coefficient between the two inductors, respectively. Notice that the magnetic coupling coefficient k is very small for this arrangement of inductors, and also that k drops off very fast for distant spirals.

3.3.3 Comparison of Different Transformer Layouts

Transformers can be realized using different configurations. These configurations offer various tradeoffs among the self inductance, the series resistance of each port, the mutual coupling coefficient, the port-to-port and the port-to-substrate capacitance, the resonant frequency, symmetry and area.

In Reference [18], a comparison of the characteristics of different transformer layouts illustrated in Figure 3.6 is given. The tapped transformer shown in Figure 3.6(a) is the most suitable for three-port applications. It relies on lateral magnetic coupling only. All winding can be implemented in the top metal layer, thereby minimizing port-to-substrate capacitance. Since the two inductors occupy separate regions, the self-inductance of each coil is maximized while the port-to-port capacitance is minimized. Also low mutual coupling ($k \simeq 0.3 - 0.5$) can be obtained because of this partial separation. Note that primary and secondary sides have asymmetric equivalent circuits.

The interleaved transformer shown in Figure 3.6(b) is suitable for four-port applications that demand for symmetry. It is realized by interwinding two identical

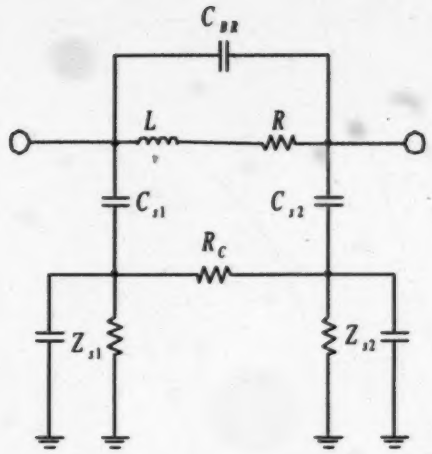


Figure 3.3: Modified spiral inductor model.

spiral inductors. Substrate capacitance can be minimized by implementation on the top level metal so that high resonant frequency may be achieved. Interleaving two inductors permits moderate coupling ($k \simeq 0.7 - 0.8$) to be realized at the cost of reduced self- inductance of the primary and secondary.

The stacked transformer shown in Figure 3.6(c) uses multiple metal layers and exploits both vertical and lateral magnetic coupling to provide the best area efficiency, the highest self-inductance and the highest coupling ($k \simeq 0.9$). This configuration is suited for both three-terminal and four-terminal utilizations. The main draw-

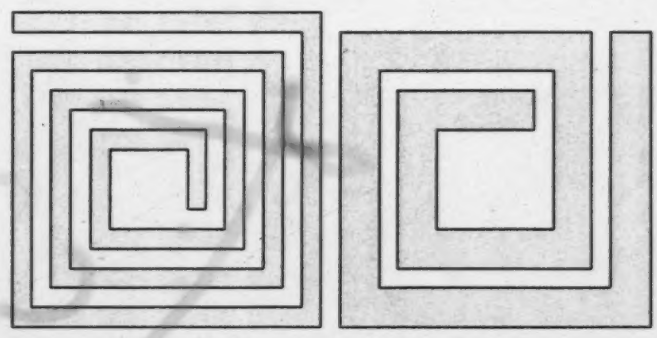


Figure 3.4: Layout of two coupled spiral structures.

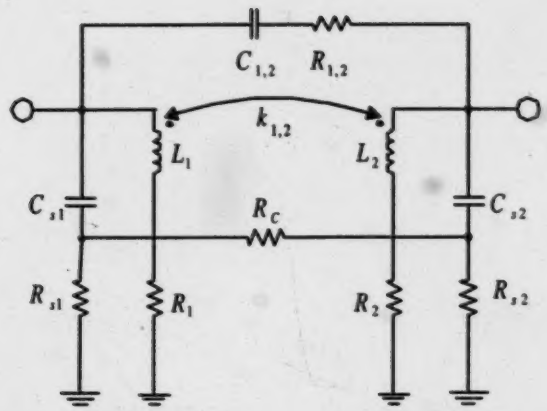


Figure 3.5: Equivalent circuit model for two coupled spirals.

back of this structure is the high port-to-port and port-to-substrate capacitance or equivalently low self-resonant frequency. One can trade reduced coupling for reduced capacitance by displacing the centers of the stacked inductors laterally or diagonally as shown in Figure 3.6(d) or Figure 3.6(e).

In order to design a more symmetric structure, one can split up the primary and secondary, and layout two halves of the primary or secondary on different metal layers. Such a structure is called metal-to-metal transformer which is shown in Figure 3.7(a) [19]. This configuration has the advantages of saving the area, offering symmetric equivalent circuits and larger k. Figure 3.7(b) represents the metal-to-metal symmetric structure more clearly with the secondary winding removed.

The comparison of different transformer layouts described above can be summarized into Table 3.1 below.

Transformer Type	Area	Coupling Coefficient	Self-Inductance	Self-Resonant Frequency	Equivalent Circuit
Tapped	High	Low	Mid	Mid	Asymmetric
Interleaved	High	Mid	Low	Low	Asymmetric
Stacked	Low	High	High	High	Asymmetric
Symmetric	Low	High	High	High	Symmetric

Table 3.1: Comparison between different transformer layouts

3.4 Design Guidelines

The design of a square-spiral inductor in silicon technologies involves a complex trade-off between the various layout and technological parameters. The number of parameters to characterize an inductor includes the outer length of the spiral A, the width of the metal tracks w, the spacing between tracks s, the number of turns N and finally the number of metal layers.

The inductance is the most accurately calculated feature. It can be evaluated by the procedures described in Greenhouse's paper [15]. The inductance is the sum of self-inductance of each segment, plus the sum of all positive mutual inductances between adjacent segments where the currents are flowing in the same direction, minus the sum of all negative inductances between segments on opposite sides of the spiral where the currents are in the opposite direction, all of which are illustrated in Figure 3.1. The total mutual inductance is much smaller than the self inductance for a single loop. It becomes dominant (a few times larger than the self inductance) when the turn count increases.

For metal lines with a rectangular cross section, the self-inductance of a line segment is [15]

$$L_{self} = 0.002l\left[ln\left(\frac{2l}{w+t}\right) + 0.50049 + \left(\frac{w+t}{3l}\right)\right] nH, \tag{3.1}$$

where l is the line length, w is the line width, and t is the thickness of the metal line. All length dimensions are in centimeters.

The mutual inductance between any two parallel segments is a function of the length of the segments and of the geometric distance between them. In general,

$$M = 2lQ, \tag{3.2}$$

where M is the mutual inductance in nH, l is the segment length in cm, and Q is the mutual inductance parameter calculated from the equation (3.3)

$$Q = ln\left[\frac{1}{GMD} + \sqrt{1+\frac{l^2}{(GMD)^2}}\right] - \sqrt{1+\frac{(GMD)^2}{l^2}} + \frac{GMD}{l}, \tag{3.3}$$

where GMD, the geometric mean distance between the two segments, is approximately

equal to the distance between the segment center. The exact value of the GMD is

$$ln(GMD) = ln(d) - \left[\frac{1}{12\left(\frac{d}{w}\right)^2} + \frac{1}{60\left(\frac{d}{w}\right)^4} + \frac{1}{168\left(\frac{d}{w}\right)^6} + \frac{1}{360\left(\frac{d}{w}\right)^8} + \frac{1}{660\left(\frac{d}{w}\right)^{10}} + ...\right], \tag{3.4}$$

where d is the center to center separation between the segments, and w is the width of the segment.

As discussed earlier this section, the total inductance of the spiral can be given as

$$L_T = L_0 + M_+ + M_-, \tag{3.5}$$

where L_T is the total inductance, L_0 is the sum of the self- inductances of all straight segments, M_+ is the sum of the positive mutual inductances and M_- is the sum of the negative mutual inductances.

The easiest parameter to set is s, the metal spacing. It should be as close as the process allows. Minimizing s can make the magnetic coupling between adjacent metal tracks maximized [20]. Tight coupling of the magnetic field can also maximize the Q-factor, increase the mutual inductance, decrease the serial resistance and reduce the total area. Minimizing s is thus mandatory.

There is an optimal value for w, the metal width. Recent study has shown that in current CMOS processes where aluminum is used, the metal width should be limited between 8μm and 13μm, depending on the other parameters but definitely smaller than 15μm before the skin effect becomes non-negligible.

The number of turns N should be restrained such that the center of the inductor is empty. Because the influence of magnetically induced losses is much more important in the inner turns of the coil, where the magnetic field reaches its maximum, inner turns should be left out.

The outer area of the inductor should be as small as possible. The diameter of the inductor directly determines the losses in the substrate. Unfortunately, parasitic electric and magnetic coupling between conductors is inversely proportional to the separation [20]. To prevent unwanted parasitic effects from disturbing the inductor's electrical characteristics, one should maintain sufficient space between the inductor and its surroundings, say, at least five metal widths ($5w$).

The inductor Q is most sensitive to the thickness and resistivity of the metal layer used in fabrication. Although the metal thickness is fixed in production technologies, metal layers in a multilevel metal process can be connected in parallel (metal stacking) with many vias to reduce the series resistance of the inductor.

The problem of designing a best-Q inductor is very complicated. It depends on many parameters that are interdependent. Concerning the technology, lower metal resistivity, thicker metal layers, lower substrate conductivity or higher substrate resistivity, thicker field oxide are desired for high quality monolithic inductors and transformers [17].

3.5 Inductor and Transformer Performance

3.5.1 Summary of the Measurements

Two square-spiral inductors and one stacked transformer were designed, fabricated, and tested in a three-metal 0.35μm CMOS technology for use in LNA designs. Approximate process parameters are listed in Table 3.2, where R_{sh} is the metal sheet resistance, t is the thickness of metal or layer, C_{sub} is the parasitic capacitance from metal to the substrate, ρ is the layer resistivity, and ϵ is the relative permittivity. Inductors were fabricated using the topmost metal 3 to minimize the series loss and to maximize the distance of the inductor from the substrate. Metal 2 is used as the "bridge layer" to make connections to the inner turn of the spiral.

Metal 3	$R_{sh} = 40\,m\Omega/\text{sq}$	$t = 0.925\,\mu m$	$C_{sub} = 7.9\,\text{aF}/\mu m^2$
Metal 2	$R_{sh} = 70\,m\Omega/\text{sq}$	$t = 0.64\,\mu m$	$C_{sub} = 12.4\,\text{aF}/\mu m^2$
Metal 1	$R_{sh} = 70\,m\Omega/\text{sq}$	$t = 0.67\,\mu m$	$C_{sub} = 30.8\,\text{aF}/\mu m^2$
Oxide Layer	$\rho = 10^{10}\,\Omega$-cm	$t = 20\,\mu m$	$\epsilon = 4$
Epitaxial Layer	$\rho = 15\,\Omega$-cm	$t = 1\,\mu m$	$\epsilon = 11.9$
Bulk Substrate	$\rho = 0.1\,\Omega$-cm	$t = 725\,\mu m$	$\epsilon = 11$

Table 3.2: 0.35μm CMOS Process Parameters

A die photo of the test layout of the inductors and transformer appears in Figure 3.8. All geometric layout parameters are given in Table 3.3. Note that layout parameters are listed for the transformer primary winding only since the secondary winding is identical.

Device	Inductance (nH)	A (μm)	w (μm)	s (μm)	N
Inductor 1	4	150	8.5	0.75	6.5
Inductor 2	10	230	9.75	1	7.5
Transformer (Primary)	9	200	8	1	6.75

Table 3.3: Layout Geometry of Inductors and Transformer

A computer simulator designed by Ali Niknejad , ASITIC [19], was used to extract the electrical parameters for a lumped-element circuit model from the inductor's specifications. Tables 3.4 and 3.5 summarize the simulation results for inductors and transformer, respectively.

Device		L (nH)	R (Ω)	C_{s1} (fF)	C_{s2} (fF)	R_{s1} (Ω)	R_{s2} (Ω)	f_{res} (GHz)
Inductor	1	4.08	11.6	128	79.2	31.3	24.7	7.07
	2	9.98	18.8	273	21.1	15.2	18.5	3.04

Table 3.4: Simulation results for the inductors

Device	L_1 (nH)	R_1 (Ω)	L_2 (nH)	R_2 (Ω)	M	k	R_G (Ω)
Transformer	9.06	25	8.99	42.7	3	0.333	-2.84

Table 3.5: Simulation results for the transformer

Here, an equivalent T-circuit model for a transformer is used for the simulation, which is different from the model for two coupled inductors shown in Figure 3.5. The T-circuit model is shown in Figure 3.9 below.

We measured these inductors and transformer, and the measurement results are posted in Tables 3.6, 3.7, 3.8, and 3.9, respectively.

3.5.2 Advice for Laying Out and Testing Inductors and Transformers

We sent several inductors and transformers to be fabricated, and we learnt that in order to make the measurement successful, some essential points should be considered before the layouts are started. First of all, it is necessary to send open pads with each design. In some cases, short or through pads are also required. These pads are used in

f (GHz)	L (nH)	R_{s1} (Ω)	R_{s2} (Ω)
0.1	4.5306	13.226	2.8481
1.0	4.6632	16.019	29.337
1.5	4.8146	19.488	45.449
2.0	4.9981	25.688	62.881

Table 3.6: Measurement results for inductor 1

f (GHz)	L (nH)	R_{s1} (Ω)	R_{s2} (Ω)
0.1	10.41	24.962	6.6572
1.0	10.95	41.188	70.109
1.5	11.79	66.523	112.27
2.0	11.634	123.96	148.59

Table 3.7: Measurement results for inductor 2

the de-embedding of inductors and transformers prior to their characterization. Second, pads should be made in metal 3 only and be consistent with input/output types. Signal input/outputs are either signal-ground (SG), ground-signal-ground (GSG), or signal-signal (SS). Input/outputs can be selected depending on the requirements of specific frequency, design, and signal count. Third, pads should match the size of available probes and meet package requirements. Figure A.18 shows a pad layout for the GSG probe with an interval of 240μm between the center of the needles. Finally, at least one ground pad is used for each probe and all ground pads must be electrically connected together to provide a low impedance and a common ground node.

f (GHz)	L (nH)	R_{s1} (Ω)	R_{s2} (Ω)
0.1	8.8934	21.078	5.582
1.0	8.6087	36.311	54.139
1.5	9.1609	56.77	86.402
2.0	9.1278	95.973	114.78

Table 3.8: Measurement results for transformer primary winding

f (GHz)	L (nH)	R_{s1} (Ω)	R_{s2} (Ω)
0.1	7.908	52.209	4.957
1.0	7.977	72.371	50.164
1.5	7.16	94.324	67.461
2.0	6.493	137.57	81.969

Table 3.9: Measurement results for transformer secondary winding

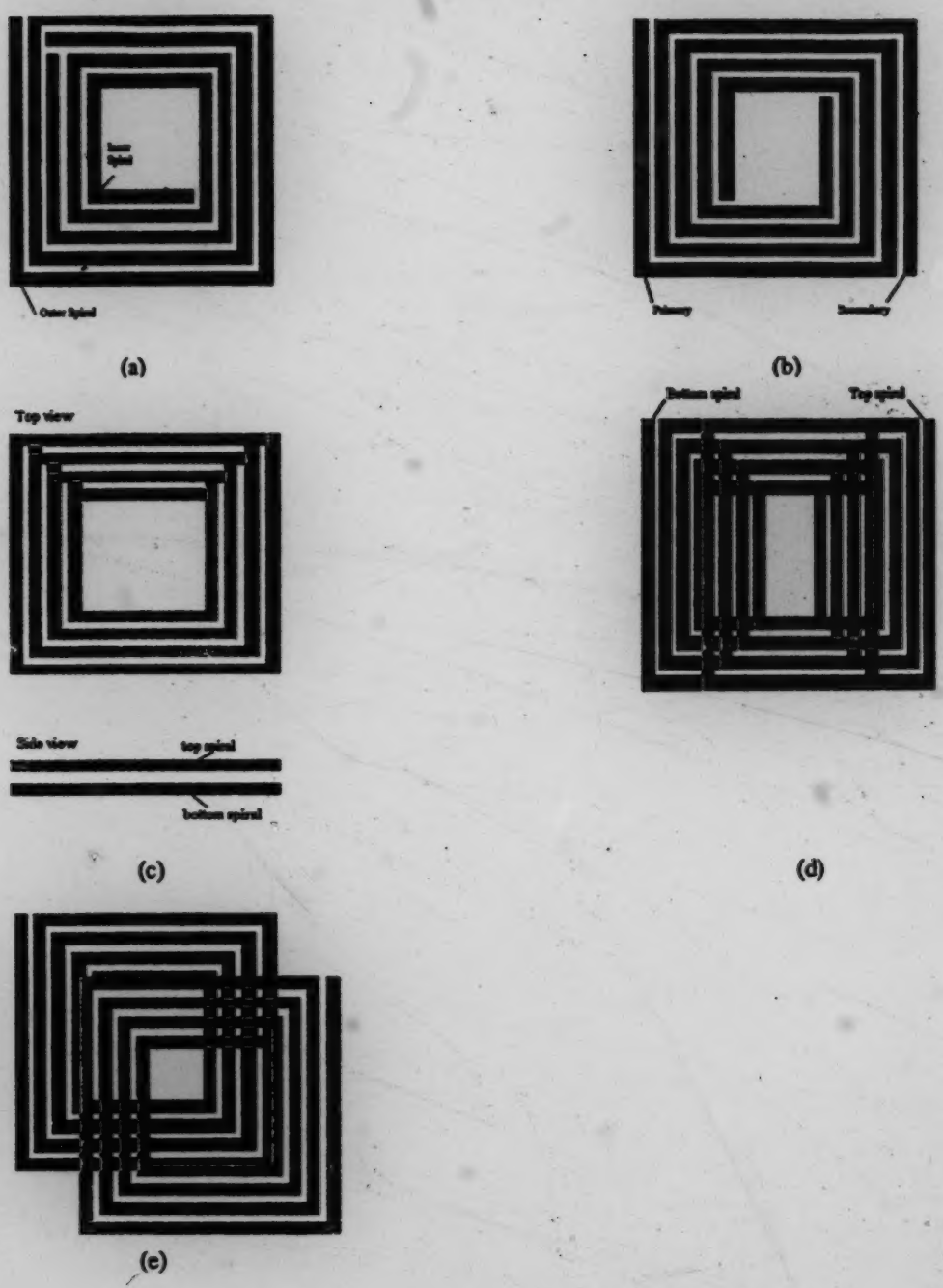


Figure 3.6: Transformer layouts I: (a) Tapped, (b) Interleaved, (c) Stacked with top spiral overlapping the bottom one, (d) Stacked with top spiral laterally shifted, (e) Stacked with top and bottom spirals diagonally shifted.

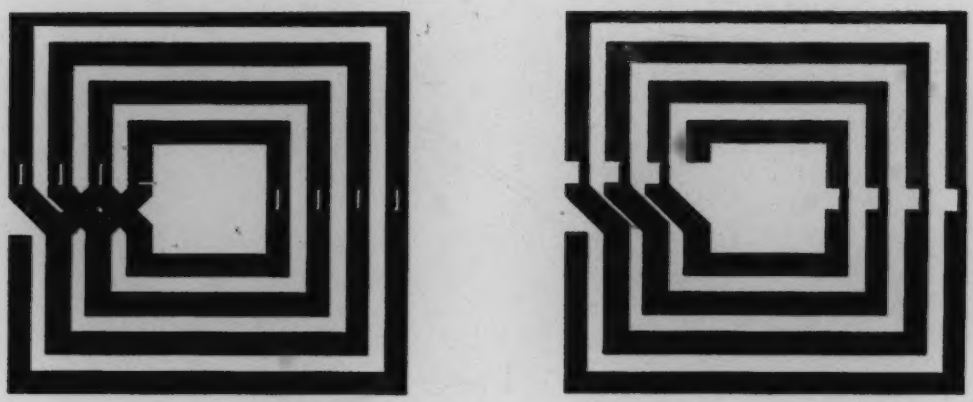


Figure 3.7: Transformer Layouts II.

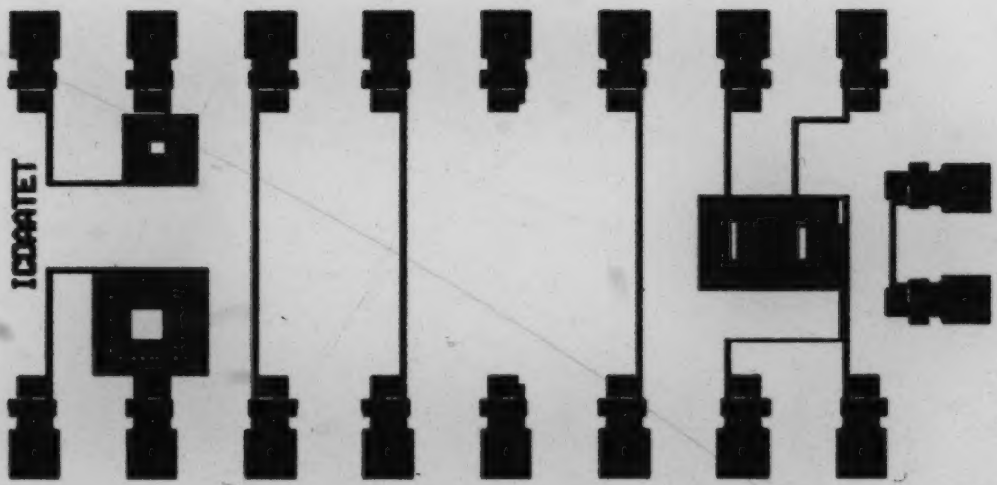


Figure 3.8: Test Layout of Inductors and Transformer.

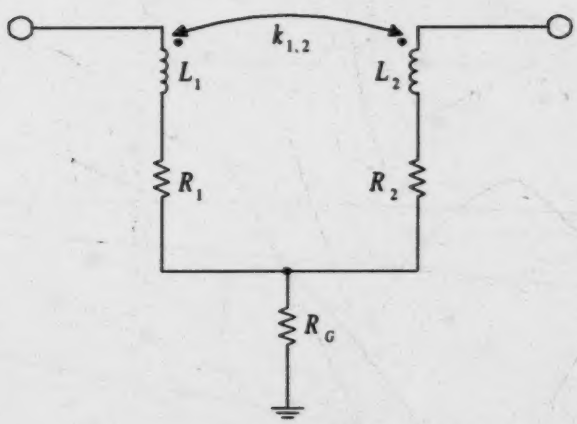


Figure 3.9: Equivalent T-circuit model for a transformer.

Chapter 4

Design Goals and Approach

4.1 Introduction

The goal of this research was to design and implement an all-CMOS RF wideband, low noise amplifier (LNA) that could be integrated into an all-CMOS receiver and eventually an all-CMOS transceiver for wireless communications applications. The majority of interest and activity for wireless communications systems is occuring at frequencies below 2 GHz, although there is increasing interest at frequencies above 2 GHz [3]. In such a high-frequency band, most LNAs are neither integrated nor CMOS. Hence this requires a different approach to LNAs.

CMOS faces some disadvantages that are less prevalent in silicon bipolar which replaced GaAs as the dominant technology for the low-GHz range. In general, the f_T of CMOS process is much lower than that of bipolar process, and as a result, CMOS is not as well-suited for high-frequency operation as bipolar. Moreover, bipolar tends to have better gain capability than CMOS due to the better mobility of electrons in bipolar. However, as mentioned earlier, CMOS does have its advantages, not the least of which are its lower cost and its ability to be integrated.

4.2 Design Goals

As stated before, the overall goal of this project was to design an RF CMOS wideband and low noise amplifier (LNA) that could be used in an integrated radio transceiver. Many of today's analog IC applications are being implemented with a power supply of 3.3 V or lower for many reasons, including power consumption and ease of integration

with digital blocks. Since CMOS technology has the ability of operating at relatively low supply voltages (e.g., 1.5-2.0 V), this LNA was designed to operate under a 2 V power supply, setting that as one of the goals of this project.

The rest of specifications of the LNA are: -3 dB bandwidth range of 2 GHz, with at least 20 dB gain and a noise figure (NF) less than 3 dB.

4.3 Approach

Once the overall requirements for the LNA have been decided, decisions on how to implement it must be made. Things such as the choice of process, the number of stages required and the types of stages have to be determined.

4.3.1 Process Choice

During the last few years, the wireless services industry has grown tremendously. With the potential for low cost, low power consumption, small dimensions and weight, and high reliability of the handset, semiconductor technology is challenged to integrate as much of the transceiver circuitry as possible and especially the RF sections.

Modern designs of wireless communications transceivers employ different technologies depending on the operating frequency of each stage. Currently, Complementary-Metal-Oxide-Semiconductor (CMOS) is the dominant technology used for integrating the baseband functions including microprocessors, DSPs, memories and digital logic. For the RF and IF parts of the transceiver, the chips are implemented using Bipolar-Complementary-Metal-Oxide-Semiconductor (BiCMOS) or bipolar processes. However, many of the external components, such as switches and power amplifiers, are still fabricated using gallium arsenide (GaAs) technology [34]. In order to obtain a small-size, low-cost, low-power system, a single-technology scheme is preferred for the maximum integration level. At this time, CMOS technology is the only solution to integrate baseband, IF and RF modules onto the same chip due to its (i) low cost, (ii) high integration levels, (iii) mixed analog/digital compatibility, (iv) low-voltage capability, and (v) production maturity [35].

For current and next generation wireless systems, the high operating frequency and steep performance requirements, conditioned under low power consumption and

cost, complicate the design of the handsets [36]. Hence, RF capable semiconductor processes face some very hard design conditions. To specify the special needs for RF operation, the performance of semiconductor processes is typically measured with respect to the active devices; that is the field-effect-transistors (FETs) [37]. A good RF capable semiconductor process should at least provide (i) an electrically/thermally insulating substrate to reduce parasitics, cross-talk, and thermal coupling, (ii) sufficient gain of active devices at the operating frequency at low bias currents, (iii) high linearity and low noise properties of active devices at low bias currents, (iv) low loss interconnects, (v) accurate active and passive devices, (vi) low cost per square millimeter, and (vii) high level of integration to minimize die size and hence cost.

With its 20-30cm wafers, CMOS technology offers the highest integration levels at the lowest cost. In RF terms, the primary performance parameters are the unity-current-gain-frequency (f_T) and the maximum-oscillation- frequency (f_{max}). Each of these parameters should be at least 10 times greater than the maximum system frequency in order to ensure sufficient overall performance [38]. Thus, for RF designs around 2 GHz, f_T and f_{max} values should exceed 20 GHz. The main component for RF CMOS design is the n-channel field-effect-transistor (NFET). For the RF performance of this active device, the most dominant achievement is the continuous down-scaling of the minimum feature size into the deep-submicron area. The minimum feature size is defined as the minimum allowed drawn gate length supported by the process technology. As the minimum obtainable channel length is reduced, yield and power performance are improved, as well as the noise performance. Most commercial processes have already reached 0.25-0.5μm, and a few prototype CMOS processes have even crossed the 0.1μm boundary. It has been verified that processes with feature size lower than 0.5μm should be able to operate in the 2 GHz range; in processes with gate lengths less than 0.1μm, f_T values can reach the 150 GHz range [39].

Although the primary objectives of increasing circuit speed and density are achieved from reduced channel length, the primary obstacles including such short-channel effects as punch-through and hot-electron degradation arise in the deep-submicron era. The different techniques are applied to scale down gate length, but they generally require scaling down of oxide thickness (t_{ox}), reducing junction depths, and increasing

the channel doping levels. All of these result in the increase of the transconductance coefficient (μC_{ox}) and parasitic capacitances [29]. Hence, according to the specifications of the LNA in this project, 0.35μm CMOS technology was chosen.

4.3.2 Multi-Stage Design

In an amplifier specification where the necessary gain cannot be obtained in single stage, the logical step is to try a multi-stage design, allowing the desired gain to be spread over multiple stages, reducing the gain for each stage to a value that can be achieved in one stage.

Using the Gain and Noise Figure equations for cascaded amplifiers discussed in Chapter One, we see that the gain of the overall LNA is simply the product of the gain of each individual stage. The equation for the total noise figure of a cascaded amplifier is a bit more complicated. The noise figure of the first stage adds directly to the overall noise figure, while the noise figure of subsequent stages are divided by the total gain of all stages before that stage. This means the first stage performance of a cascaded amplifier is very critical to the overall noise figure. Moreover, as discussed in previous chapters, there is a tradeoff between gain and bandwidth. With the increase of the number of stages, the overall bandwidth decreases while increasing gain. Due to these restrictions on the wideband LNA design, one must realize that there is some optimum point, and the best thing is not to have a large number of stages, and only a minimal gain per stage. In this design, the optimum was a two-stage amplifier. The first stage was designed to achieve both input matching and low noise operation while the second stage is employed to provide good output matching.

Now that the number of stages have been determined, the topology of the various stages must be determined. Circuit topology refers to all portions of the circuit: the active devices, the biasing network and the input/output matching networks. In the next chapter, the specific circuit-level implementation will be addressed.

Chapter 5

Implementation

Once the architectural decisions have been made, the wideband low noise amplifier must be realized in a circuit form. The desired -3 dB bandwidth should be realized while maintaining as high a gain and low a noise as possible.

5.1 Circuit Design

5.1.1 Transistor Biasing

The biasing of the active device is a critical factor in the overall functionality of the design. A good reference is insensitive to temperature, process, and supply voltage variations. In CMOS integrated circuit design, the reference voltages can be derived from the power supplies using resistor and the MOSFET [40]. Figure 5.1 shows the basic idea. The voltage divider shown in Figure 5.1(a) is formed with two resistors. It has the advantage of simplicity, temperature insensitivity, and process insensitivity, that is, changes in the sheet resistance have no effect on the voltage division. The main problem with this circuit is that in order to reduce the power dissipation (i.e., the current through the resistors), the resistors must be made large. Since large resistors require a large area on the die, this voltage divider may not be practical in many cases. In the following, the behavior of the last two circuits in Figure 5.1 will be analyzed.

The Resistor-MOSFET Divider

The resistor-MOSFET voltage divider shown in Figure 5.1(b) is formed with a resistor and an n-MOSFET. The reference voltage generated in this divider is equal to the V_{gs}

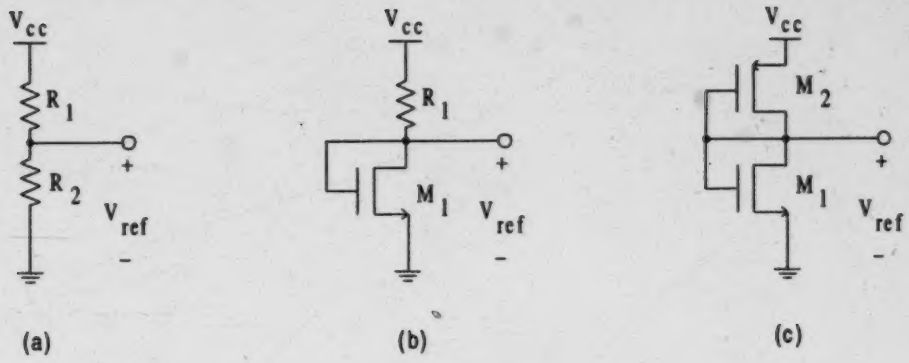


Figure 5.1: Implementation of voltage dividers in CMOS.

of the n-MOSFET. We know that when an n-MOSFET is operated with its channel pinched off, that is, $V_{ds} \geq V_{gs} - V_t$ and $V_{gs} \geq V_t$, it is operating in the saturation region. We can write

$$\begin{aligned} I_D &= \frac{\beta_1}{2}(V_{gs} - V_t)^2 \\ &= \frac{V_{cc} - V_{ref}}{R_1}, \end{aligned} \tag{5.1}$$

or the reference voltage is given by

$$V_{ref} = V_t + \sqrt{\frac{2I_D}{\beta_1}}, \tag{5.2}$$

where

$$\beta_1 = \mu_n C_{ox} \cdot \frac{W_1}{L_1}. \tag{5.3}$$

A modification of this basic divider is shown in Figure 5.2.

The MOSFET-only Divider

The MOSFET-only voltage divider shown in Figure 5.1(c) generates a reference voltage equal to the voltage on the gates of the MOSFETs with respect to ground. Since $I_{D1} = I_{D2}$, we can establish the following equation:

$$\frac{\beta_1}{2}(V_{ref} - V_{ss} - V_{tn})^2 = \frac{\beta_2}{2}(V_{cc} - V_{ref} - |V_{tp}|)^2.$$

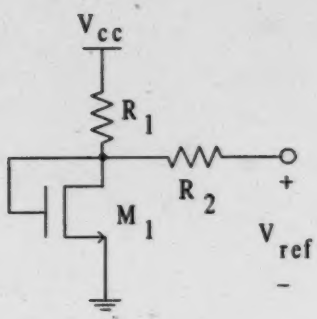


Figure 5.2: Modification of the resistor-MOSFET voltage divider.

From here, the reference voltage is given by

$$V_{ref} = \frac{V_{cc} - |V_{tp}| + \sqrt{\frac{\beta_1}{\beta_2}}(V_{ss} + V_{tn})}{\sqrt{\frac{\beta_1}{\beta_2}} + 1},$$

or

$$\frac{\beta_1}{\beta_2} = \left[\frac{V_{cc} - V_{ref} - |V_{tp}|}{V_{ref} - V_{ss} - V_{tn}}\right]^2,$$

if the reference voltage and the power supply are known. Notice that, here MOSFET-only voltage divider is used between V_{cc} and V_{ss}, which means the reference voltage is equal to the voltage on the gates of the MOSFETs with respect to V_{ss}, instead of ground.

In this project, the biasing circuit configurations are chosen as Figure 5.1(b) or Figure 5.2 due to their simplicity and small layout size.

5.1.2 Input Stage

The input stage must be able to handle the input signal without contributing significantly to the total distortion. In addition, its noise must be low since the noise performance of the entire amplifier is determined mainly by the noise of the input stage. Furthermore, the low noise and low distortion must be achieved without excessive power consumption [41], and 50 Ω matching should be achieved at the input as well.

The topology of the input stage for the LNA is shown in Figure 5.3, formed by transistors T_3 and T_4. The input stage is cascoded for a number of reasons. The first

is to minimize the effect of the Miller capacitance by forcing no amplification at the output node of the lower device T_3's gate-to-drain overlap capacitance C_{gd}. Specifically, the Miller effect tends to substantially lower the input impedance, complicating the task of matching to the input [42]. In addition to diminishing the Miller effect, the use of the cascode gives better gain from its input to its output because of an increase in the output resistance of the device. Furthermore, the cascode provides good LNA's reverse isolation between its output and its input. In a standard, single MOS device amplification stage, the output signal can be fed back to the input by the gate-to-drain capacitance C_{gd}. At high frequencies, the impedance of this capacitance decreases, and more and more of the output will be fed back to the input. The cascode, on the other hand, has no such a connection, and thus provides larger amount of reverse isolation. Actually, the only way of feedback to the input from the final stage is through parasitic interconnect capacitance, which can be minimized through good layout.

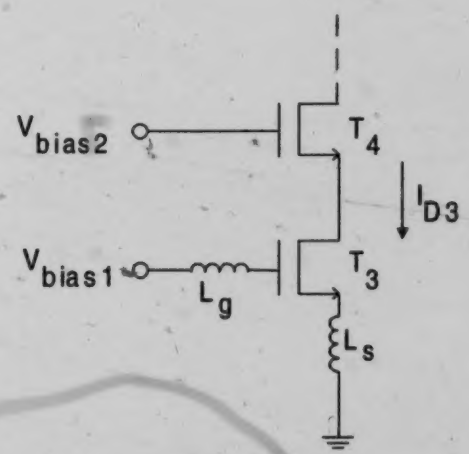


Figure 5.3: Input stage of the LNA.

As shown in Figure 5.3, the LNA must present the proper input impedance to achieve input matching. For this purpose, inductive degeneration is employed in the source of T_3. This degeneration produces a real term in the LNA's input impedance that is used in matching to the source resistor $R_s = 50\ \Omega$. The impedance matching

network is formed by inductors L_s and L_g along with the gate-to-drain capacitance C_{gs} of T_3. As discussed in Chapter 2, under matching conditions to R_s,

$$\omega_0^2 C_{gs}(L_g + L_s) \approx 1, \tag{5.4}$$

and

$$R_s = L_s \frac{C_{gs}}{g_m}. \tag{5.5}$$

It has been known that the quality factor Q of the matching network is also the voltage gain between the LNA input and the gate of T_3 [43]. As shown by equations (5.6) and (5.7),

$$Q = \frac{1}{2 g_m \omega_0 L_s}, \tag{5.6}$$

and

$$\begin{aligned} NF &= 1 + \gamma \frac{1}{Q^2 \cdot g_m R_s} \\ &= 1 + \gamma \frac{2\omega_0 L_s}{Q \cdot R_s}, \end{aligned} \tag{5.7}$$

where γ is a bias-dependent coefficient whose value can be taken as 2/3 for long-channel devices, or 2-3 for short-channel devices. A higher Q requires lower g_m from T_3 and reduces the NF due to T_3's drain current noise. However, the use of high Q is limited [43], since a high Q network results in large reactance levels associated with C_{gs} and L_g, compared to 50 Ω. In this LNA design, Q is limited to 2-3, which is beneficial for reducing both current consumption and NF.

5.1.3 Output Stage

The output stage is shown in Figure 5.4. It consists of a bridging capacitor C_B and a "T-coil" network which is the combination of three inductors (L_1, L_2, and L_3). The use of the T-coil is for bandwidth extension. The combination of three inductors can be realized conveniently as a pair of magnetically coupled inductors, namely, a transformer which is shown in Figure 5.5(a) with controlled coupling. An equivalent circuit model of such a transformer is shown in Figure 5.5(b).

From Figure 5.5(a), we can establish the following set of equations:

$$\dot{V}_1 = j\omega L_1 \cdot \dot{I}_1 - j\omega M \cdot \dot{I}_2, \tag{5.8}$$

$$\dot{V}_2 = -j\omega M \cdot \dot{I}_1 + j\omega L_2 \cdot \dot{I}_2. \tag{5.9}$$

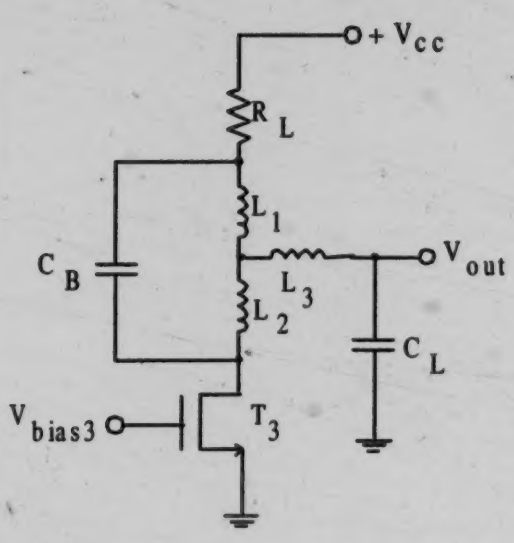


Figure 5.4: Output stage of the LNA.

Equations (5.8) and (5.9) can be written as

$$\begin{aligned} \dot{V}_1 &= (j\omega L_1 + j\omega M)\cdot \dot{I}_1 - j\omega M \cdot \dot{I}_1 - j\omega M \cdot \dot{I}_2, \\ \dot{V}_2 &= -j\omega M \cdot \dot{I}_1 - j\omega M \cdot \dot{I}_2 + j\omega (L_2 + M)\cdot \dot{I}_2. \end{aligned}$$

On rearranging the two equations above,

$$\dot{V}_1 = j\omega (L_1 + M)\cdot \dot{I}_1 - j\omega M \cdot (\dot{I}_1 + \dot{I}_2), \tag{5.10}$$

$$\dot{V}_2 = -j\omega M \cdot (\dot{I}_1 + \dot{I}_2) + j\omega (L_2 + M)\cdot \dot{I}_2. \tag{5.11}$$

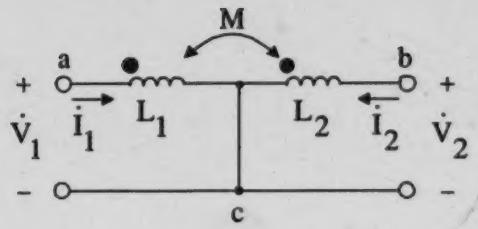
Equations (5.10) and (5.11) are the same set of equations as that for the circuit in Figure 5.5(b), which is thus proven as the equivalent circuit for Figure 5.5(a). With the coupling as shown,

$$L = L_{ab} = (L_1 + M) + (L_2 + M) = L_1 + L_2 + 2M, \tag{5.12}$$

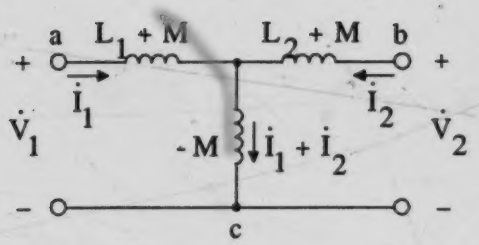
$$L_{ac} = (L_1 + M) - M = L_1, \tag{5.13}$$

$$L_{bc} = (L_2 + M) - M = L_2, \tag{5.14}$$

where the addition of mutual inductance $-M$ in the equivalent circuit produces the correct self-inductance. The coupling coefficient is always positive, and can be defined



(a)



(b)

Figure 5.5: (a) Coupled inductors forming a T-coil, (b) Equivalent circuit with mutual inductance M.

as

$$k = |\frac{M}{\sqrt{L_1 L_2}}|. \tag{5.15}$$

The application of the T-coil for bandwidth enhancement has resulted in general form of the bridged T-coil circuit depicted in Figure 5.6. The coil is terminated in R_L, but the load C_L is connected to the center-tap. We can see that at DC, the input impedance Z_{in} of the T-coil is R_L. At high frequencies, due to the bridging capacitor C_B, it is also R_L. The derivation of the input impedance is given in [52]. Let us assume that $L_1 = L_2 = L_1$, and introduce additional limitations:

$$C_B = \frac{C_L}{4}(\frac{1-k}{1+k}), \tag{5.16}$$

$$L = R_L^2 C_L, \tag{5.17}$$

$$L_1 = L_2 = \frac{L}{2(k+1)}. \tag{5.18}$$

Here $k = M/L_1$ as usual. For $Z_L = 1/sC_L$, the transfer function for the load is

$$\frac{V_o}{I_i}(s) = \frac{R_L}{1 + \frac{1}{2}R_L C_L s + \frac{1}{4}(\frac{1-k}{1+k})R_L^2 C_L^2 s^2}. \tag{5.19}$$

The derivation of the transfer function is rather difficult and is given in [52] as well.

The form of the transfer function is the same as that for the series peaking circuit but with twice the speed improvement. For maximally flat amplitude (MFA), namely, the Butterworth-type response, the transfer function is expressed as

$$\frac{V_o}{I_i} = \frac{R_L}{1 + \sqrt{2}P + P^2}. \tag{5.20}$$

Comparing the denominator of equation (5.20) with that of equation (5.19), we can see that

$$\frac{1}{4}(\frac{1-k}{1+k})R_L^2 C_L^2 s^2 = P^2,$$

and

$$\frac{1}{2}R_L C_L s = \sqrt{2}P.$$

Solving the above two equations, we get the coupling coefficient

$$k = \frac{1}{3},$$

and

$$P^2 = \frac{1}{8}R_L^2 C_L^2 s^2. \tag{5.21}$$

As we know, the -3 dB bandwidth is found in the usual way by setting the magnitude of equation (5.20) to $1/\sqrt{2}$. Assuming $P = j\Omega$, then we can set up the following equation

$$\frac{1}{(1-\Omega^2)^2 + 2\Omega^2} = \frac{1}{2}.$$

Substituting the result of $\Omega = 1$ into equation (5.21), we obtain that

$$\frac{1}{\sqrt{8}}R_L C_L \omega_{-3dB} = 1.$$

Then

$$\omega_{-3dB} = 2.83\frac{1}{R_L C_L}. \tag{5.22}$$

This is the greatest bandwidth improvement, nearly three times that of the original RC circuit. It is shown that a coupling coefficient of 1/3 yields an MFA response,

while a k of 1/2 leads to maximally flat group delay (MFED) if the same analysis is applied. These coupling coefficients are not too large, thus they are not hard to implement in practice.

As we discussed in Chapter 3, to save the die area, a transformer can be implemented as a pair of spiral inductors that have been placed on top of each other and offset appropriately to obtain the desired amount of coupling. If a small bridging capacitance is added across the inductors to create a parallel resonance, a further improvement can be achieved. The increased circulating currents associated with the resonance help to push the bandwidth out even further [9].

5.1.4 Theoretical Calculations

Prior to the calculations, some useful parameters and results for a MOSFET in the typical 0.35μm technology will be summarized. In this project, only n-channel MOSFETs are considered and utilized.

The threshold voltage for an n-channel MOSFET, V_t, is approximately 0.6V. The dielectric constant of the gate oxide, ε_{ox}, is $0.35\times10^{-13}F/m$. The gate oxide thickness, t_{ox}, equals $0.73\times10^{-6}m$. Since the gate-oxide capacitance per unit area, C_{ox}, can be expressed as

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}.$$

Substituting ε_{ox} and t_{ox} into the above equation, we can obtain that C_{ox} is equal to $4.7945\times10^{-7}F/cm^2$. The transconductance parameter for an n-channel MOSFET is given by

$$KP_n = \mu_n \cdot C_{ox}, \tag{5.23}$$

where μ_n is the mobility of the holes in an NMOS transistor, whose typical value in the 0.35μm CMOS process is 424 $cm^2/V \cdot sec$. The transconductance parameter, KP_n, is calculated by substituting μ_n and C_{ox} into equation (5.23), which results in a value of 203.3 $\mu A/V^2$. Furthermore, another transconductance parameter for an n-channel MOSFET, β_n, is defined as

$$\beta_n = KP_n \cdot \frac{W}{L}. \tag{5.24}$$

As discussed earlier, the drain current (I_D) for the n-channel MOSFET operating in the saturation region can be calculated using equation (5.1), which can be

represented as

$$I_D = \frac{KP_n}{2} \cdot \frac{W}{L} \cdot (V_{gs} - V_t)^2 = \frac{\beta_n}{2}(V_{gs} - V_t)^2, \quad \text{for } V_{ds} \geq V_{gs} - V_t \text{ and } V_{gs} \geq V_t. \tag{5.25}$$

Also, the transconductance g_m can be written as

$$g_m = \sqrt{2I_D}\sqrt{\mu_n C_{ox} W/L} = \sqrt{2I_D \cdot \beta_n}. \tag{5.26}$$

The gate-source capacitance of the MOSFET in the saturation region, C_{gs}, is given by

$$C_{gs} = \frac{2}{3} W L C_{ox}; \tag{5.27}$$

and the unity gain frequency, ω_T, of the n-channel MOSFET is

$$\omega_T = \frac{g_m}{C_{gs}}. \tag{5.28}$$

Calculations for Biasing Circuits

In this project, the transistors biasing circuit for the cascoded input stage of the LNA is shown in Figure 5.7, which is the combination of Figure 5.1(b) and Figure 5.2. Assuming that $V_{cc} = 2V$, $V_{bias1} = 0.8V$ and $V_{bias2} = 1.5V$, when $R_2 \ll R_1$, it can be easily found that $V_{bias1} \simeq V_{gs}$ and $I_1 \simeq I_D$. From equations (5.1), (5.2) and (5.3), a new equation can be obtained:

$$(V_{bias1} - V_t)^2 = 2 \cdot \frac{V_{cc} - V_{bias1}}{R_1 \cdot KP_n} \cdot \frac{L_1}{W_1}. \tag{5.29}$$

Then

$$R_1 \cdot W_1 = \frac{2 \times (2 - 0.8)V}{(0.8 - 0.6)^2 V^2} \cdot \frac{0.35\mu m}{203.3\mu A/V^2} = 1.034 \times 10^5 \Omega \cdot \mu m,$$

where R_1 is in Ω and W_1 is in μm. Assuming that R_1 is set to 400 Ω, then from the above equation, W_1 can be calculated to be 258.5 μm.

Similarly, the product of $R_3 \cdot W_2$ can be found the same way. From the following equation:

$$(V_{bias2} - V_t)^2 = 2 \cdot \frac{V_{cc} - V_{bias2}}{R_3 \cdot KP_n} \cdot \frac{L_2}{W_2},$$

$R_3 \cdot W_2$ can be calculated as 2128.57 $\Omega \cdot \mu m$. If R_3 is chosen as $2K\Omega$, then W_2 is equal to $1.06\mu m$.

Calculations for Input Circuit

The input stage of the wideband LNA is shown in Figure 5.3. Assuming $I_D = 5mA$, from equation (5.25), we can obtain that

$$W_3 = \frac{2 \cdot I_D \cdot L_3}{KP_n \cdot (V_{bias1} - V_t)^2}.$$

Substituting values into the above equation, we get

$$W_3 = \frac{2 \times 5 \times 10^{-3}A \times 0.35\mu m}{203 \times 10^{-6}A/V^2 \times (0.8 - 0.6)^2 V^2} = 431\mu m.$$

The transistor T_3's transconductance g_{m3} is given by equation (5.26), then it can be calculated as

$$\begin{aligned} g_{m3} &= \sqrt{2I_D \cdot KP_n \cdot W_3/L_3} \\ &= \sqrt{2 \times 5 \times 10^{-3}A \times 203 \times 10^{-6}A/V^2 \times 431\mu m/0.35\mu m} \\ &= 0.05A/V. \end{aligned}$$

Using equation(5.27), C_{gs} can be found as

$$C_{gs} = \frac{2}{3} \times 431\mu m \times 0.35\mu m \times 4.7945F/cm^2 = 0.482pF.$$

In order to match the input impedance to source resistance ($R_s = 50\Omega$), a set of matching conditions should be satisfied. From equations (5.4) and (5.5), we can obtain that

$$L_s = R_s \cdot C_{gs}/g_m = 50 \times 0.482 \times 10^{-12}/0.05 = 0.482nH.$$

If $Lg \gg L_s$ and the operating resonant frequency f_0 is set to 1.9 GHz, then

$$\begin{aligned} L_g &\approx \frac{1}{\omega_0{}^2 C_{gs}} \\ &\approx \frac{1}{(2\pi \times 1.9 \times 10^9)^2 \times 0.482 \times 10^{-12}} \\ &\approx 14.56nH. \end{aligned}$$

As unit gain frequency ω_T can be expressed as equation (5.28), substituting the values of g_m and C_{gs} into it, we can obtain that

$$f_T = \frac{1}{2\pi} \cdot \frac{g_m}{C_{gs}} = \frac{0.05}{2\pi \times 0.482 \times 10^{-12}} = 16.51GHz.$$

From equation (5.6), the quality factor Q of the matching network can be found

$$\begin{aligned} Q &= \frac{1}{2g_m\omega_0 L_s} \\ &= \frac{1}{2 \times 0.05 \times 2\pi \times 1.9 \times 10^9 \times 0.482 \times 10^{-9}} \\ &= 1.738. \end{aligned}$$

The noise figure NF can be obtained from equation (5.7). For long-channel devices, $\gamma = 2/3$; the best-case of NF can be calculated as

$$\begin{aligned} NF_{best} &= 1 + \gamma \frac{1}{Q^2 \cdot g_m R_s} \\ &= 1 + \frac{2}{3} \times \frac{1}{1.738^2 \times 0.05 \times 50} \\ &= 1.09. \end{aligned}$$

While the worst-case of NF can be obtained, when $\gamma = 3$ for short-channel devices. The calculation can be expressed as follows:

$$NF_{worst} = 1 + 3 \times \frac{1}{1.738^2 \times 0.05 \times 50} = 1.4.$$

Calculations for Output Circuit

In Figure 5.4, the output stage of the wideband LNA is depicted. If the 10 dB voltage gain $(A_v)_{dB}$ is required and the load resistor (R_L) is set to 200 Ω, then from the following equation:

$$A_v = g_{m4} \cdot R_L, \tag{5.30}$$

the transconductance g_{m4} can be obtained as:

$$g_{m4} = \frac{10^{10/20}}{200\Omega} = 0.016S.$$

As shown in Figure 5.3, we can easily find that $I_{D3} = I_{D4}$ and $L_3 = L_4$, then from equation (5.26), g_m is obviously proportional to $\sqrt{W}$. Hence, W_4 can be set up as

$$W_4 = \left(\frac{g_{m4}}{g_{m3}}\right)^2 \cdot W_3.$$

Substituting the given values into the above equation, the width of the transistor T_4 is equal to

$$W_4 = \left(\frac{0.016}{0.05}\right)^2 \times 431\mu m = 44.13\mu m.$$

As stated in the early part of this chapter, the bandwidth can be almost tripled if the coupling coefficient (k) of the T-coil in the output stage equals 1/3 which leads to a Butterworth-type (maximally flat amplitude) response. Then let us choose the load capacitor with the capacitance of $C_L = 600fF$ and $k = 1/3$. From the output stage equations (5.15), (5.16), (5.17) and (5.18), the bridging capacitor C_B, the primary or secondary inductance L_1 or L_2 ($L_1 = L_2$) and the mutual inductance M which are the components of the "T-coil" network, can be calculated as follows:

$$\begin{aligned} C_B &= \frac{600fF}{4}\frac{1-1/3}{1+1/3} = 75fF, \\ L &= 200^2 \times 600 \times 10^{-15} = 24nH, \\ L_1 &= L_2 = \frac{24}{2 \times (1+1/3)} = 9nH, \\ M &= \frac{1}{3} \times 9 = 3nH. \end{aligned}$$

As -3 dB bandwidth can be obtained from equation (5.22). For the given R_L and C_L, we can get that

$$\omega_{-3dB} = 2.83 \times \frac{1}{200 \times 600 \times 10^{-15}} = 2.358 \times 10^{10} rad/sec.$$

5.2 Wideband LNA Layout

The circuit design process is a trial and error process with many attempts; the completion of the circuit design sets up the stage for the next important step, namely, the layout of the LNA. Since the layout adds considerable capacitance to each node, especially the sensitive nodes in the signal path, care must be taken to minimize resistive losses across metal lines, and to ensure that other losses in the circuit do not diminish the overall performance of the circuit achieved in the previous design.

5.2.1 Active Device Layout

The first important point to realize is that the transistors must be drawn in such a fashion that the RC time constant of the physical gate does not adversely impact the signal drive. Typically, transistors in analog circuits are much wider than transistors in digital circuits. In practice, active devices with widths of several hundred or even thousand μm are possible. The devices used in this design are large, therefore

they cannot just be laid out as one physically long transistor. The input signal has frequency near 1.9 GHz, and so if the physical resistance of the poly-silicon gate in series with the gate capacitance of the device has a very slow time constant, the signal driving the transistor itself will be interrupted. The simplest way to solve this problem is to use a comb structure, where a wide transistor is composed of smaller transistors connected in parallel.

A simplified layout of this approach is shown in Figure 5.8(a), where four transistors that have a common gate are connected in parallel. Figure 5.8(b) shows the circuit corresponding to the layout in Figure 5.8(a), where the transistors have been drawn in the same relative positions. Figure 5.8(c) shows the same circuit redrawn differently, where it is clear that the circuit consists of four transistors connected in parallel. These four MOSFETs operate as a single MOSFET with a width equal to the sum of the individual MOSFET's widths, assuming equal lengths. Notice that node 2 has a much greater junction capacitance than node 1, because node 2 has a larger total junction area and especially a larger perimeter. Thus, when the equivalent transistor is connected to a circuit, node 1 should be connected to the more critical node. Also notice that the large number of contacts in wide junction regions are used to minimize the contact impedance, and thus to greatly minimize voltage drops that would occur due to the relatively high resistivity of silicon junctions compared to the resistivity of the metal that overlays the junctions and connects them. With this layout, two benefits are achieved: (1) smaller layout size and (2) reduction of source and drain depletion capacitances.

Practically, the maximum width of a single finger of the comb MOSFET device structure is limited to 200μm in 0.35μm CMOS process. As we know, the value of f_{max} for MOSFET can be expressed as follows [44]:

$$f_{max} = \frac{f_T}{2\sqrt{2\pi f_T R_g C_{gd} + G_{ds} R_{in}}}, \tag{5.31}$$

where G_{ds} is the output conductance, R_{in} is the input resistance consisting of the gate, source and channel components, and R_g is the distributed gate resistance. The gate resistance, R_g, is a dominant parameter governing f_{max} which is simply given by [45]

$$R_g = \frac{R_{\square} W}{3n^2 L}, \tag{5.32}$$

where $R_{\square}$ is the poly-silicon sheet resistance, L is the gate length, W is the total gate width and n is the number of gate fingers.

From equations (5.31) and (5.32), we find that for fixed unit finger width W_u, two times increase of n produces two-fold R_g reduction and two-fold C_{gd} and G_{ds} increase, thus resulting in no variation of f_{max}. While for fixed total width W, two times decrease of W_u leads to four-fold R_g reduction, in turn $\sqrt{2}$ times increase of f_{max}. The value of f_{max} increases with reducing unit finger width W_u, but the increase rate of f_{max} decreases for 5μm of W_u because the term of $G_{ds}R_{in}$ in equation (5.31) becomes dominant. Therefore, W_u of 10-15μm is the optimal width range to obtain the good RF performance for 0.35μm poly-silicon gate MOSFETs. For example, the size of 360μm/0.35μm transistor in this design is split into 36 smaller transistors connected in parallel, with each having the 10μm width. The layout of this transistor is shown in Figure A.1.

5.2.2 Capacitor Layout

As passive components, capacitors play an important role in MOS technology. Many MOS technologies that are used to implement analog functions have two layers of poly-silicon. The additional layer provides an efficient capacitor structure and an extra layer of interconnection. The plan view of a typical poly capacitor structure is shown in Figure 5.9. The plate separation is comparable to the gate oxide thickness (t_{ox}) of the MOS transistors. Usually, the bottom plate of the capacitor will be the first layer of poly-silicon. The interconnection of the top plate, which is also poly-silicon, can often be done in the first-level metal with contacts to the poly-silicon plate. Ideally, capacitance is given by

$$C_1 = \frac{\varepsilon_{ox}}{t_{ox}} A_1 = C_{ox} X_1 Y_1, \tag{5.33}$$

where ε_{ox} is the electric constant of the gate oxide, C_{ox} is the gate oxide capacitance per unit area, and A_1, X_1 and Y_1 are the area, length and width of the capacitor, respectively.

An important aspect of the capacitor structure is the parasitic capacitance associated with each plate. The largest parasitic capacitance is formed by the bottom plate and any conductors beneath it, which could be either the substrate or a well diffusion

in CMOS whose terminal is electrically isolated [6]. This bottom-plate parasitic is proportional to the bottom-plate area and its typical value is as large as 10 to 30 percent of the main capacitance and often severely limits circuit performance.

The top-plate parasitic is caused by the interconnect metalization or poly-silicon which connects the top plate to the rest of the circuit, plus the parasitic of the transistor to which it is connected.

In this project, capacitors are taken square, where the side-length X_1 is equal to the side-width Y_1. From equation (5.33), the ideal capacitor size can be derived as

$$A_1 = X_1 Y_1 = X_1^2 = \frac{C_1}{C_{ox}}, \tag{5.34}$$

where C_{ox}, the gate oxide capacitance per unit area, is approximately 889.9 $aF/\mu m^2$ for the CMOS 0.35μm process. For example, to build up a 90fF poly-poly capacitor, the ideal side-length of the square will be chosen to be 10.06μm, according to equation (5.34). Actually, error exists between the true capacitor size and the ideal capacitor size. Normally, the size of the ideal capacitor is larger than that of the practical one. The major sources of the error in realizing capacitors are due to over-etching which causes the area to be smaller than the area of the layout masks, and an oxide-thickness gradient across the surface of the microcircuit. The former effect is usually dominant and can be minimized by placing a number of smaller, unit-sized capacitors in parallel with the overall capacitance being the desired capacitor, similar to what is usually done for transistors. The latter effect can be minimized by interspersing the unit-sized capacitors in a common-centroid layout so the gradient changes affect both capacitors in the same way. Since oxide-thickness variations are not large in a small area, this common-centroid layout is not used. In this project, these two ways to minimize the capacitor error are not applied, because the capacitors are not too large. The layout of the poly-poly capacitor with 90fF capacitance is illustrated in Figure A.2.

5.2.3 Resistor Layout

Integrated resistors can be realized using different conductors. A popular choice is poly-silicon, which is a deposited and etched material and is more resistive than metal. However, these days most poly-silicon is silicided specifically to reduce re-

sistance. Other choices include diffused or ion-implanted regions such as junctions, wells, or base regions. Another possibility is deposited and etched thin-film resistors. Among all these resistor materials, poly-silicon resistors have a reasonably low parasitic capacitance per unit area, the lowest voltage coefficient and a moderate temperature coefficient. These are the major reasons why poly-silicon resistors are preferred over other types of resistors even though they often require more area due to the low resistivity.

Regardless of the type of resistor used, the total resistance equation is given by

$$R = \frac{L}{W} R_{\square}, \tag{5.35}$$

where L is the length of the resistor, W is the width of the resistor, and $R_{\square}$ is the resistance per square. Usually, $R_{\square}$ can be expressed as

$$R_{\square} = \frac{\rho}{t}, \tag{5.36}$$

where $\rho = 1/(q\mu_n N_D)$ is the resistivity, t is the thickness of the conductor, and N_D is the concentration of carriers, which we assume are electrons. Since poly-silicon resistors are usually doped by p^+, the concentration of holes, the expression of the resistivity can be rewritten as $\rho = 1/(q\mu_p p^+)$, where μ_p is the hole mobility.

Usually, the nominal resistivity of most poly-silicon layers that are utilized in standard CMOS processes is in the vicinity of 5 to $10\Omega/\square$ [9], so poly-silicon is mainly for moderately small-valued resistors. In order to obtain medium-sized resistors, a serpentine layout is used which is similar to that shown in Figure 5.10. When calculating the total resistance of such a structure, we must take the bends and the end contacts into consideration. For example, each bend contributes 1.1μm length and each finger contributes 7.65μm length in this project. For the structure shown, a total of 10 fingers and 9 bends contribute 76.5μm and 9.9 μm length, respectively. Thus, the total length of the resistor is 86.4μm. If the resistivity of the resistor is $8.1\Omega/\square$ for 0.35μm CMOS technology, according to equation (5.35), then the total resistance is

$$R = \frac{86.4}{0.35}\square \times 8.1\Omega/\square \simeq 2000\Omega. \tag{5.37}$$

The layout of the poly-silicon resistor with resistance of 2000Ω is shown in Figure

A.3. Notice that an additional impedance due to the contacts should be added to the total resistance.

5.2.4 Final Circuit Schematic and Layout

The final circuit schematic, along with all coupling capacitors and inductors as well as the DC bias at each node, is shown in Figure A.4 with detailed component values and Figure A.5 with detailed DC operating points. The final layout is shown in Figure A.6; note that parts A and B are shown in detail in Figures A.7 and A.8. This circuit which we refer to as LNA1 is designed for the voltage gain with 10 dB; only single cascoded stage is utilized. The finished LNA1 appears cleanly as shown in Figure 5.11.

While another kind of final full circuit schematic designed for the voltage gain with 20 dB is illustrated in Figure A.9 with detailed component values and Figure A.10 with detailed DC operating points. In order to obtain the necessary gain, two cascaded stages are applied to this circuit (LNA2) with the first stage implemented using the cascoded design. The final layout is shown in Figure A.11 with details of parts A and B given in Figures A.12 and A.13. The theoretical calculation procedure is very similar to that in the single stage. Here it is omitted. The finished LNA2 appears cleanly as shown in Figure 5.12.

Notice that, the feedback capacitor C_{F1} with the capacitance of 90fF in LNA1 and C_{F2} with the capacitance of $50fF$ in LNA2 are both used to prevent the circuit from overshooting and ringing. Furthermore, each bond wire and connection pin on the package accumulate approximately 5nH of inductance. It should be taken into account in creating the schematics. At frequency beyond 1 GHz, this inductance dominates the input and output impedances and creates noticeable negative feedback between stages of the amplifier. In order to reduce this negative feedback, it is necessary to separate the power and ground bond pads used in different stages. The pads shown in the layouts are specially for RF applications. The detailed pad layout for GSG signal is shown in Figure A.18.

The component values and transistor widths for the designs are given in Table 5.1. The total layout area for LNA1 (see Figure A.6) is 1000μm $\times$ 1500μm, and 1500μm $\times$ 830μm for LNA2 (see Figure A.11).

Parameters	LNA1	LNA2
M_1 (μm)	200	200
M_2 (μm)	1.8	6
M_3 (μm)	90	150
M_4 (μm)	360	375
M_5 (μm)	-	150
R_1 (Ω)	400	525
R_2 (Ω)	2000	2000
R_3 (Ω)	-	225
R_L (Ω)	200	100
L_g (nH)	4	10
L_1 (nH)	9	9
L_2 (nH)	9	9
M (nH)	3	3
C_L (fF)	50	50
C_B (fF)	50	50
C_F (fF)	90	50

Table 5.1: Design circuit parameters

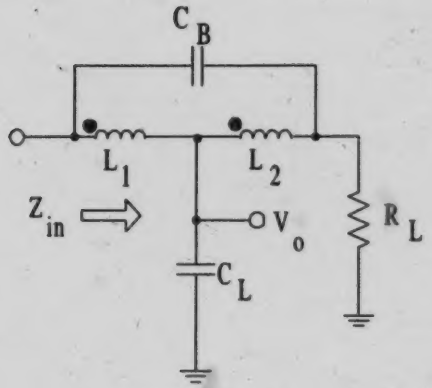
5.3 Simulation Results

Transistor level simulations for two schematics illustrated in Figures 5.11 and 5.12 are performed using the Analog Artist, Composer workflow and the SpectreS circuit simulator from Cadence 0.35μm CMOS technology. The wideband LNAs are both to be driven by 100μV AC signal. Figures A.14 (a) and A.14 (b) respectively show the output voltage and the input impedance versus frequency, which correspond to the one-stage wideband LNA schematic. The output impedance characteristics for the same schematic is shown in Figure A.15 as well. Similarly, the corresponding voltage at the output of the second stage, the input impedance and the output impedance versus frequency for the two-stage wideband LNA are shown in Figures A.16 (a), A.16 (b) and A.17, respectively.

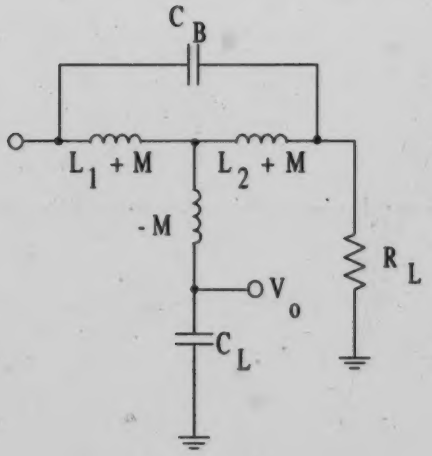
The simulation results for two LNAs are summarized in Table 5.2. Here, I_1, I_2 and I_3 are stage currents. All the results almost meet the design requirements. One point needs to be mentioned is that both designs achieve the input impedance matching when the frequency is below 1 GHz. But for the output impedance case, it is worse.

Parameters	LNA1	LNA2
I_1 (mA)	3.246	2.772
I_2 (mA)	5.037	3.844
I_3 (mA)	-	9.759
I_{total} (mA)	8.283	16.37
V_{out} (μV)	329.2	731.7
Gain (dB)	10.35	17.29
-3dB BW (MHz)	2038.24	1941.48
Power consumption (mW)	16.57	32.74
Z_{in} @1 GHz	55.33	59.96
Z_{in} @2 GHz	80.72	85.26
Z_{out} @1 GHz	97.96	91.95
Z_{out} @2 GHz	105.24	151.62

Table 5.2: Summary of simulation results for two LNAs ($V_{cc} = 2V$; $V_{in} = 100\mu V$)



(a)



(b)

Figure 5.6: (a) Bridged T-coil with C_L load, (b) Equivalent circuit.

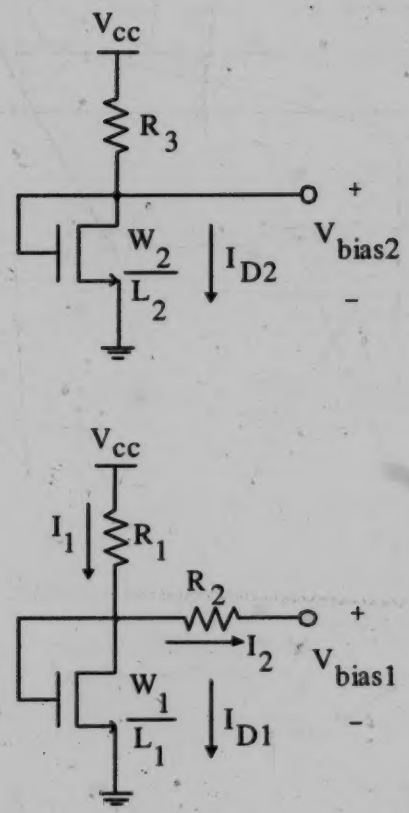


Figure 5.7: Biasing circuit for the cascoded input stage of the LNA.

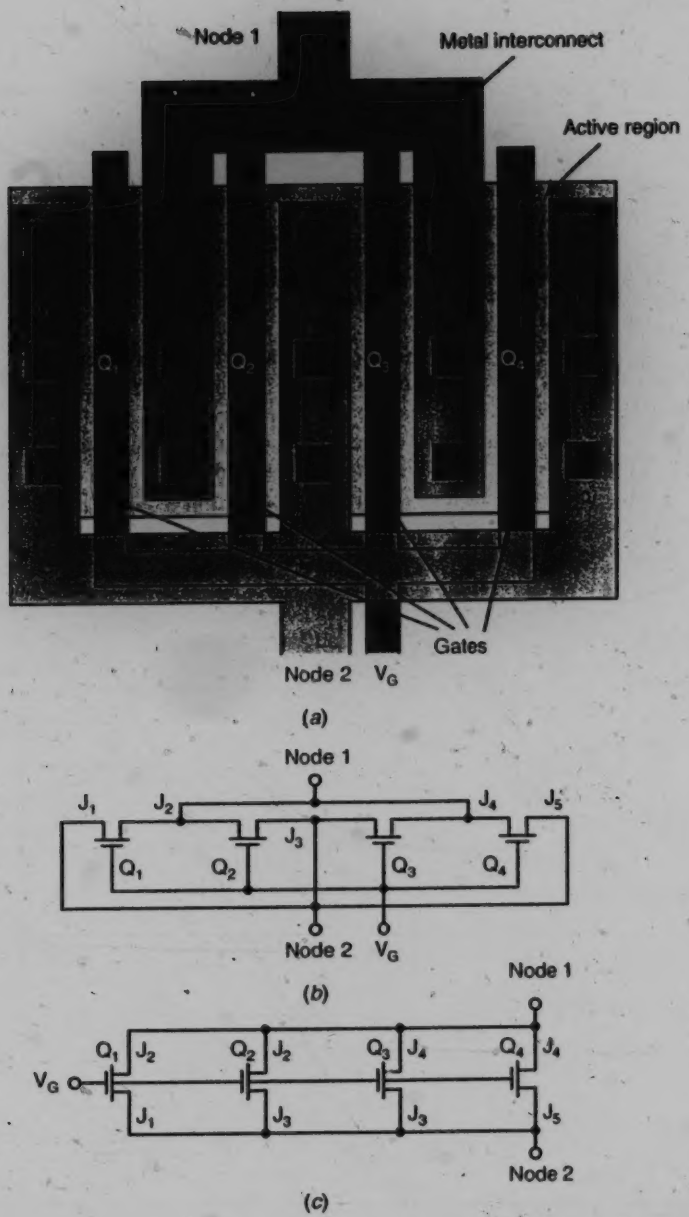


Figure 5.8: Connecting four transistors in parallel to realize a single large transistor: (a) the layout, (b) the schematic drawn in the same relative positions as the layout, and (c) the circuit redrawn to make the parallel transistors more obvious.

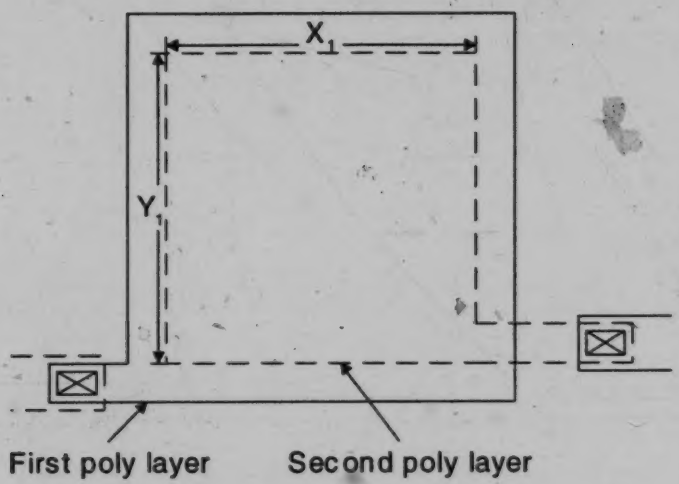


Figure 5.9: Plan view of a typical poly-poly capacitor.

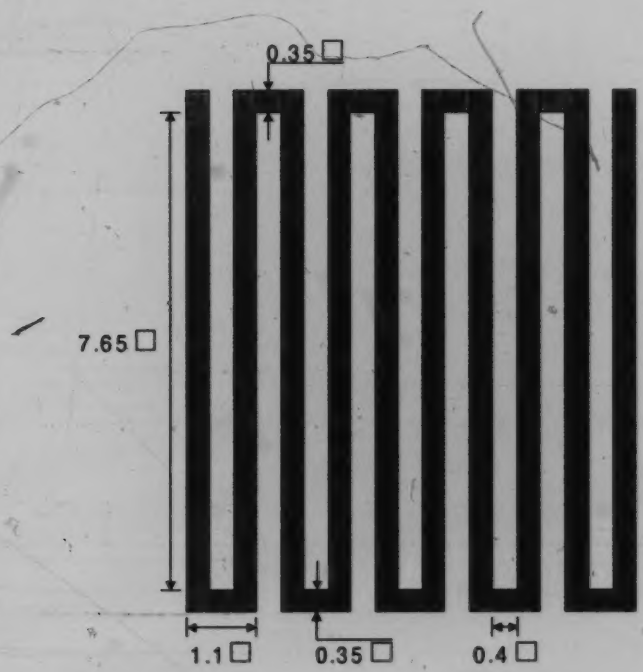


Figure 5.10: A typical layout for an integrated poly-silicon resistor.

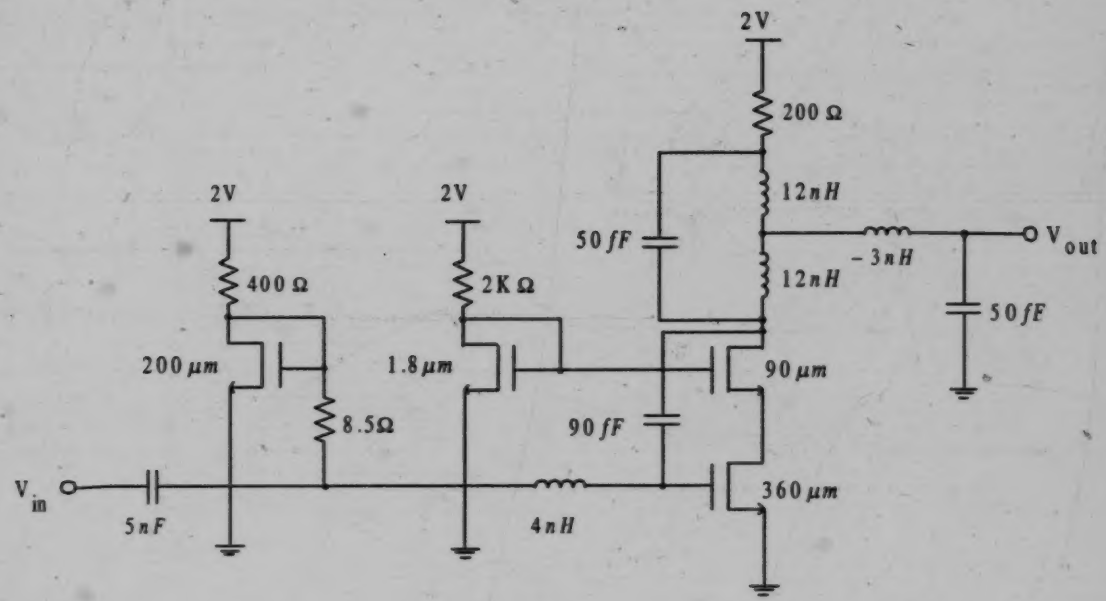


Figure 5.11: Complete LNA1.

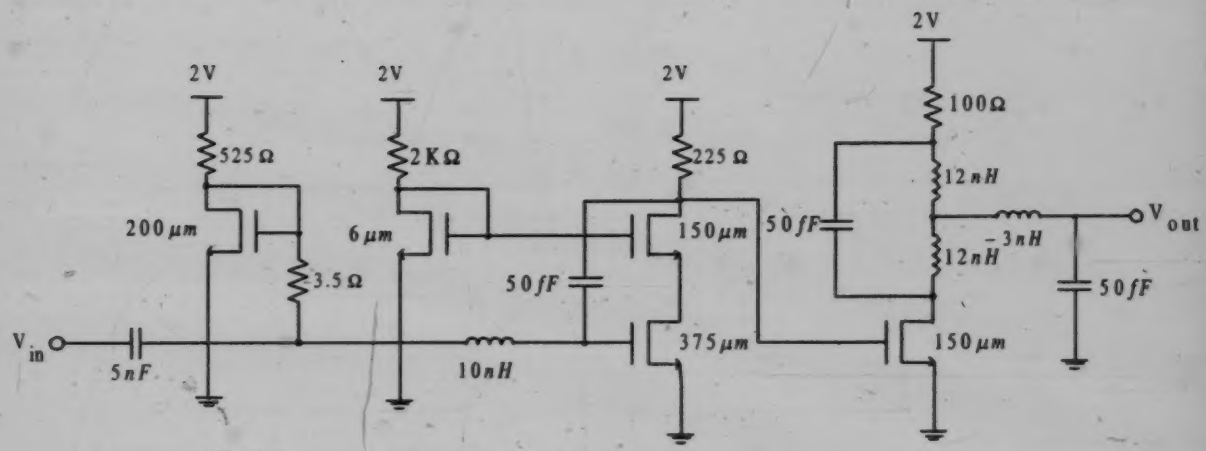


Figure 5.12: Complete LNA2.

Chapter 6

Experimental Results

Both wideband LNAs are fabricated in a three-metal, 0.35μm CMOS process offered by Taiwan Semiconductor Manufacturing Company Ltd. (TSMC).

6.1 Overview

The general setup for electronic measurements consists of three elements: the device under test (DUT), the measuring probes and the measuring instruments; this is shown in Figure 6.1. Such a setup is valid for both low frequency and high frequency measurements. However, the components in the setup differ depending on the range of frequencies at which the measurements will be done. At low frequencies, the measuring instruments are mainly oscilloscopes and digital multimeters, while at high frequencies, network analyzer and spectrum analyzer become the key measuring instruments. Also, for low-frequency tests, the probes could be any kind of electrical wires. For RF and microwave tests, measurements can be done using either wafer probes, coaxial cables or SMA connectors. For RFIC measurements, two methods can be applied. One is to use PCB test fixtures; the other is on-wafer measurement. Since measurement on-Silicon-wafer can be completed without having to bond them into fixture [47], it is more direct, immediate and accurate. Therefore, the on-wafer measurement method is chosen in this project.

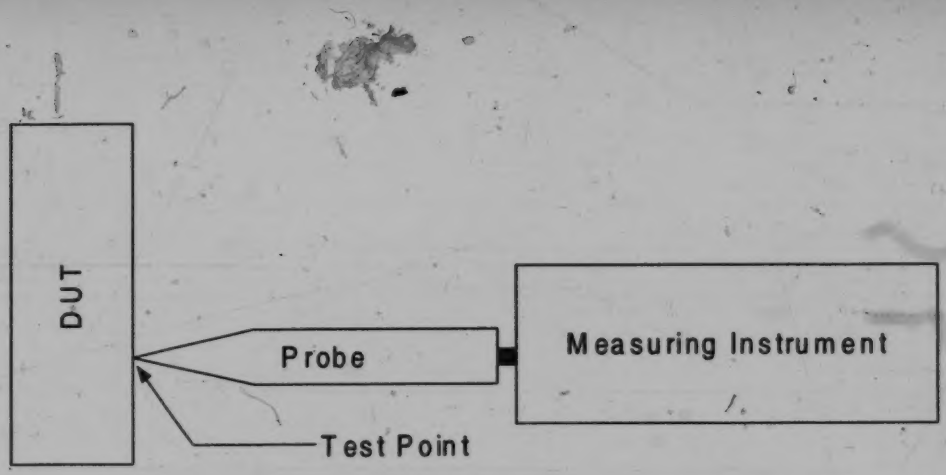


Figure 6.1: General setup for electronic measurements.

6.2 Test Equipment

The measuring instruments used in the testing are listed in this section. All DC voltages were measured with a Fluke Model 75 Digital Multimeter (Accuracy: ± 0.5% ± 1 Digit). All amplifier currents were measured with a Fluke Model 75 Digital Multimeter (Accuracy: ± 2% ± 1 Digits). The S-parameters were measured with the HP 85046A S-Parameter Test Set (300 KHz - 3.0 GHz), and the HP 8753A Network Analyzer (300 KHz - 3.0 GHz) which was calibrated by GGB Industries Calibration Substrate CS-8 Kit. A Brunelle Model 1040 Power Meter was used to supply 2-V DC voltage. In addition, HP 8495B Attenuator, Cambridge Instruments StereoZoom 6 Photo microscope, GGB Industries Picoprobe station, and Newport 460A Series for signal-ground (SG) and ground-signal (GS) wafer probes were used as well. All measurement results were plotted with a HP 7475A Plotter. All measurements were made with both input and output terminated in 50 Ω SMA loads.

6.3 Chip Testing

The final step in this project is to test the chips. All measurements were performed using bare dice fixed on a piece of glass by vacuum suction and probed by the GS and SG wafer probes. The wafer probes were connected to a probe station that allow the movement of the probes in the x, y and z direction. The other end of the probes were then connected to the measuring instruments through SMA connectors

and 50 Ω coaxial cables. The microscope was used to help make proper contact of the spring loaded tips of the probes with the surfaces of the RF pads. As we know, accuracy in network analysis is greatly influenced by factors external to the network analyzer. Parts of the measurement setup such as interconnecting cables and test sets all introduce variations in magnitude and phase that can mask the actual performance of the DUT [48]. Hence, the calibration step before the DUT is actually measured is necessary. Here the Model CS-8 calibration substrate, a standard to which the measurement system (network analyzer + cabling + probe) can be connected, is used for calibrating out the unavoidable errors and losses in the network analyzer, its associated cabling and the probe for the on-wafer testing [49]. Three major calibration methods have been developed to calibrate network analyzer systems: SOLT (Short, Open, Load, Through), LRM (Line, Reflect, Match), and LRL (Line, Reflect, Line). The SOLT method is the most commonly used which was applied for this project.

Figure 6.2 illustrates the frequency response experimental setup. Here the probe station described above is not shown in figure. The S-parameters of the chip can be measured using the HP 8753A network analyzer. Note that the network analyzer requires a signal level at the R input in the range of 0 to -35 dBm to phase lock the internal source [48]. Therefore, a signal of -10 dBm was input to the chip. An external attenuator with attenuation of 10 dB was used here. The nominal supply voltage of the fabricated amplifier (LNA1) was 2 V. The measurements were performed by a wafer probing system. The pads connected at the output of the LNA1 facilitate direct measurement of the performance of the wideband LNA.

The wideband frequency response calibration with 10 dB attenuation is shown in Figure 6.3. Figure 6.4 is the input reflection coefficient, S_{11}, of the wideband LNA1. A return loss of -14.7 dB is usually adequate to obtain a satisfactory response. Figure 6.5 shows the voltage gain of the LNA1 at 2 V. The amplifier provides a forward gain S_{21} of 11.64 dB, swept from 300 KHz to 400 MHz. While the -3 dB bandwidth ranges from 300 KHz to 642 MHz. The measured reverse gain S_{12} and output reflection coefficient S_{22} are shown in Figures 6.6 and 6.7, respectively.

An overview of the measured parameters of the fabricated LNA1 is given in Table 6.1, showing well defined gain, low output reflection coefficient and excellent reverse isolation. Notice that the measured -3 dB bandwidth is quite different from that

expected. After checking the final layout, the reason is found: the capacitive load was forgotten to be drawn in the layout, while it was included in the schematic.

LNA1	Measured Parameters
Supply voltage (V)	2
Total current (mA)	7.21
Power dissipation (mW)	14.42
-3 dB Bandwidth (MHz)	0.3 - 642
Forward gain S_{21} (dB)	11.64
Reverse isolation S_{12} (dB)	<-32
Input reflection coefficient S_{11} (dB)	-14.7
Output reflection coefficient S_{22} (dB)	-12.7

Table 6.1: Summary of LNA1 measurements

A die photo of the experimental chip (LNA1) is shown in Figure B.1. The active area is 1×1.5 mm^2. Figures B.2 is the photo for the inductor with inductance of 10nH; and Figure B.3 is the photo for the stacked transformer with the primary and secondary winding of 9nH each and coupling coefficient $k = 1/3$.

The measurements of LNA2 follows the same procedure as that of LNA1. It seems that the S_{21}, the gain from input to output, is negative. In fact, it is a kind of attenuation rather than the gain through the chip. The reason for this is determined after reviewing the layout and testing the DC operating voltages of the NMOS transistors. From the LNA2 layout, we can see clearly that metal traces in power nets are quite narrow, which are connected to the power supply pads. Therefore, the traces will be damaged when the current goes too high. Furthermore, the input DC voltage, namely V_{gs} of transistor M_3, is much larger and the output DC voltage, namely V_{ds} of transistor M_5 in the final stage, is much smaller than the desired values, respectively. This may cause violation in the transistor working equation ($V_{gs} - V_{ds} \leq V_t$). As a result, the transistors would not work properly. In addition, the capacitor as the load is still missing.

Another big issue for on-wafer S-parameter measurements is de-embedding the parasitics of external interconnects and pads. When performing high-frequency measurements on wafer, a two-step correction procedure has to be followed [50]. First, the measurement system has to be calibrated, defining a reference plane for the S-

parameter measurements at the probe tips using a standard calibration technique (SOLT) which is described in the earlier part of this chapter. Second, the on-wafer parasitics have to be characterized, so that from the measurement the actual parameters can be obtained. Usually, a ceramic impedance-standard-substrate (ISS) containing high precision calibration standards is used for on-wafer measurements [51]. Provided the substrate containing the DUT has similar loss and coupling characteristics as the ISS, this type of calibration will offer a calibrated reference plane close to the probe tips. Since high-loss silicon CMOS substrate can not be easily represented by a stand-alone ISS, an alternative is to implement a calibration kit which resides on the same wafer as the DUT. By this approach, measurement pads and external interconnects can be included in the calibration and thereby de-embedded. Among five different de-embedding methods (open pad, open/short, open/short/ thru, two-port, and 3-step de-embedding), only open/short de-embedding (OSD) was applied in the LNA1 design. As we know, interconnects in CMOS technology are typically realized using aluminum traces. Since the ISS uses gold for metalization, this means that the contact impedance increases when the probes are lifted from the ISS onto the CMOS chip. This additional series loss can be determined by adding a short standard to the set. The detailed analysis and the steps for the de-embedding realization will not be addressed here.

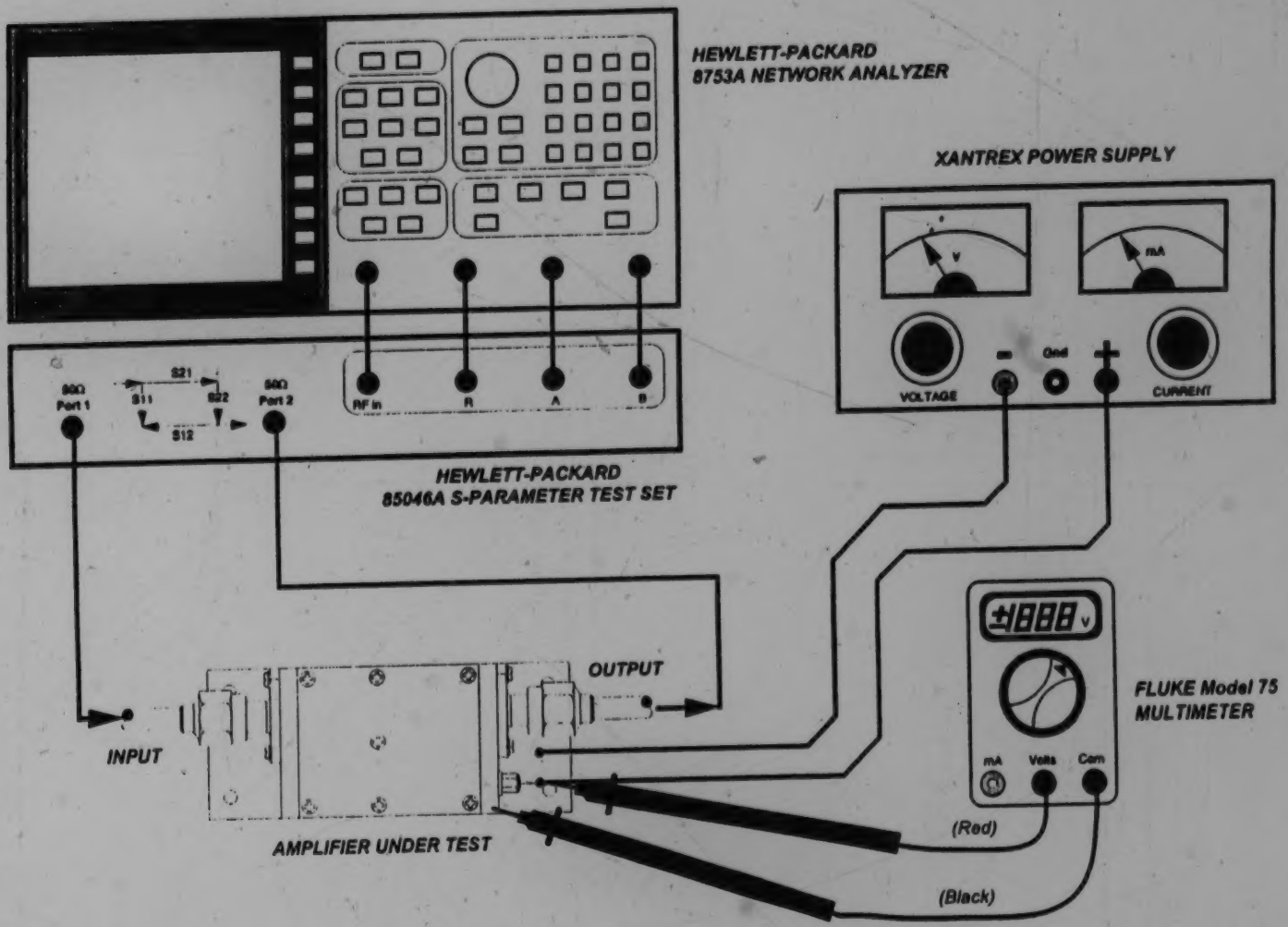


Figure 6.2: Frequency response experimental setup.

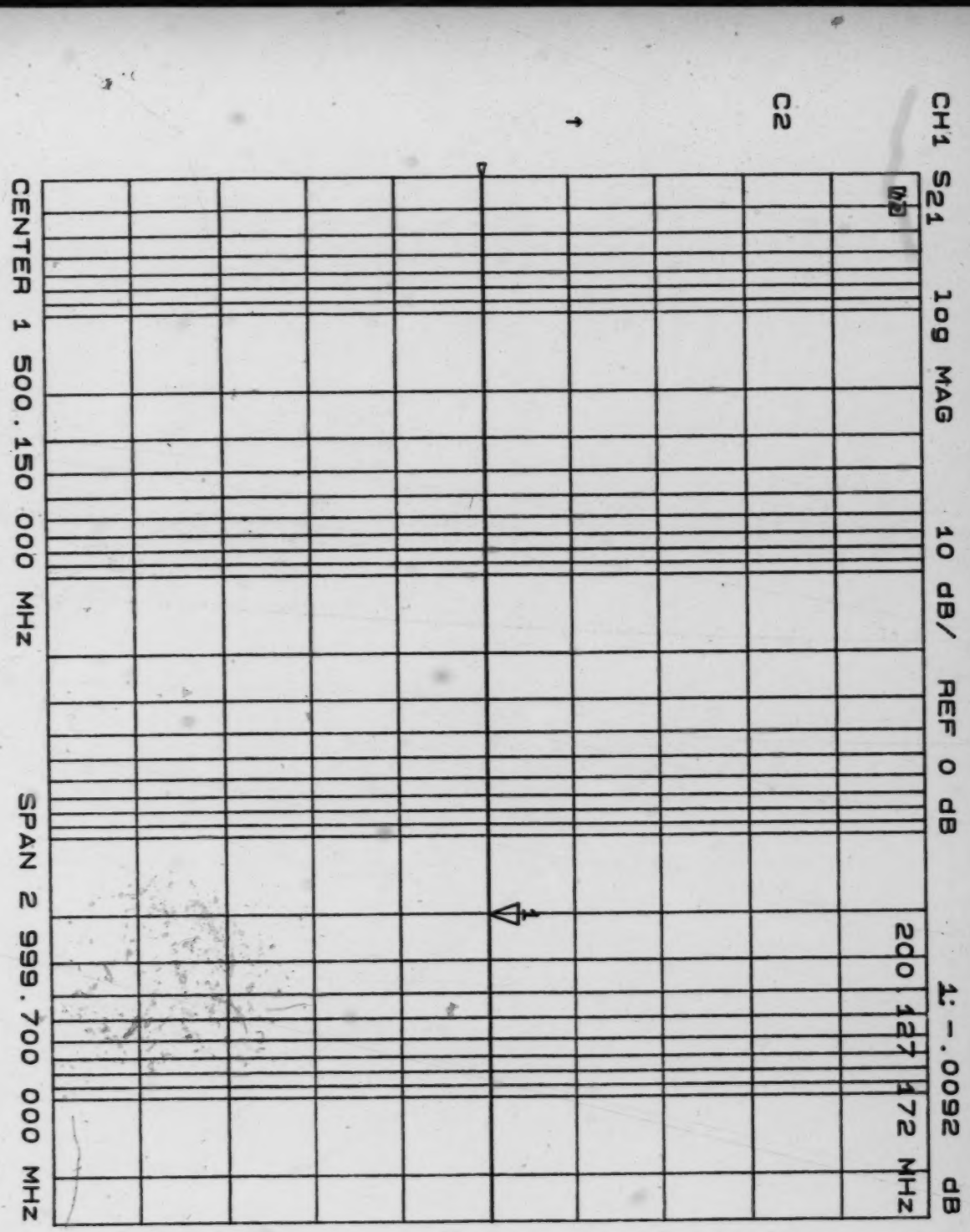


Figure 6.3: Wideband frequency response calibration with 10 dB attenuation.

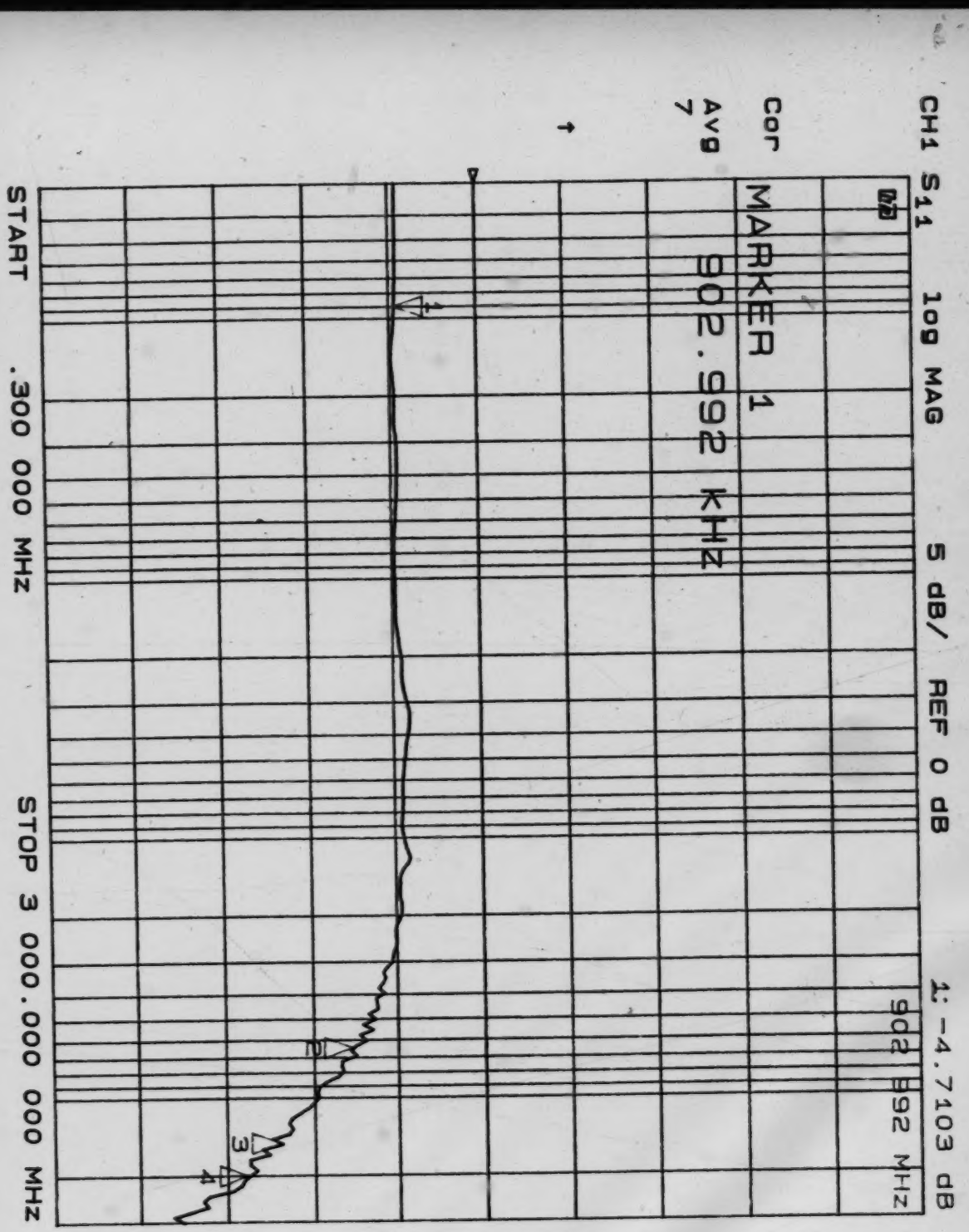


Figure 6.4: S_{11} vs. frequency for LNA1.

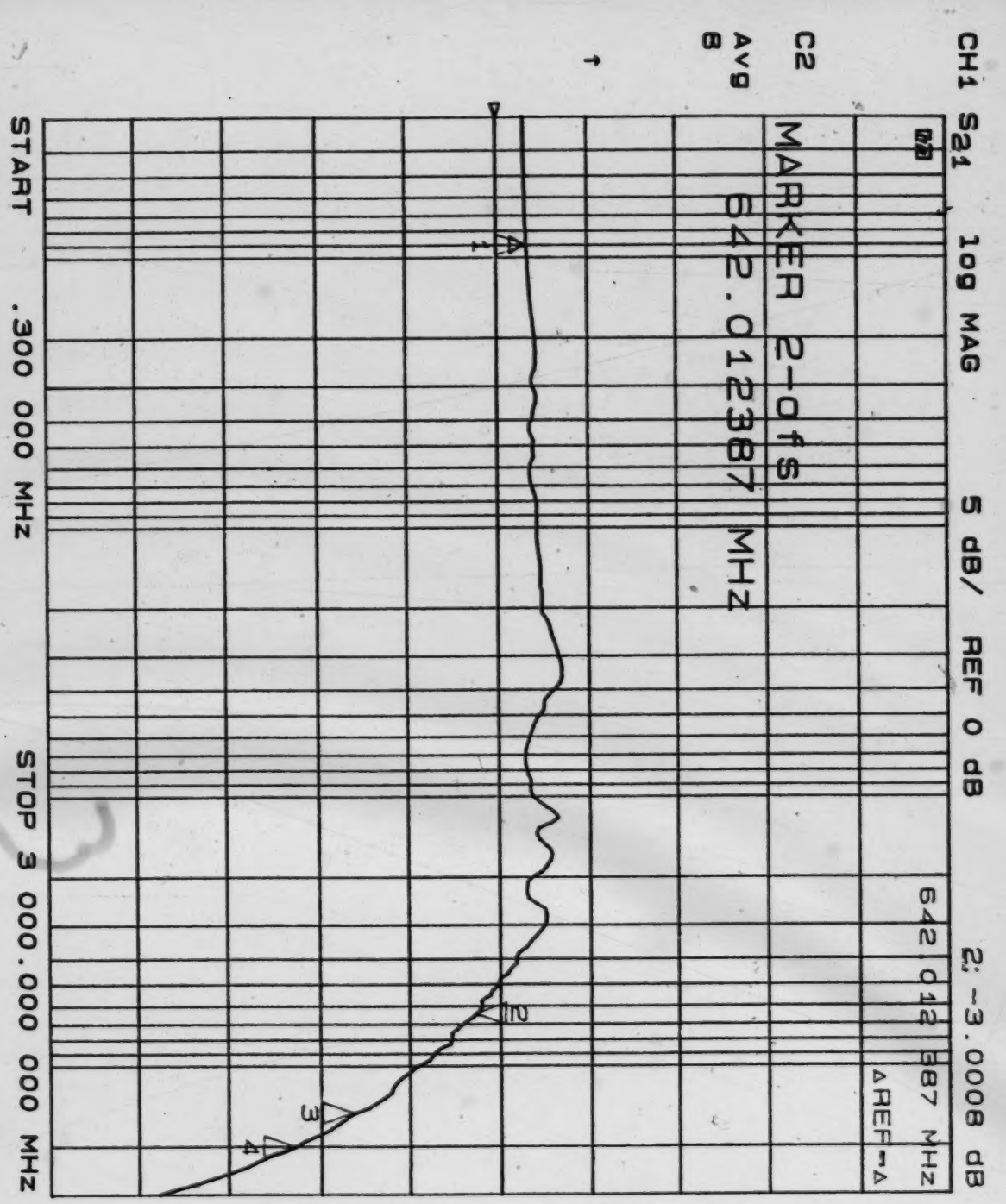


Figure 6.5: S_{21} vs. frequency for LNA1.

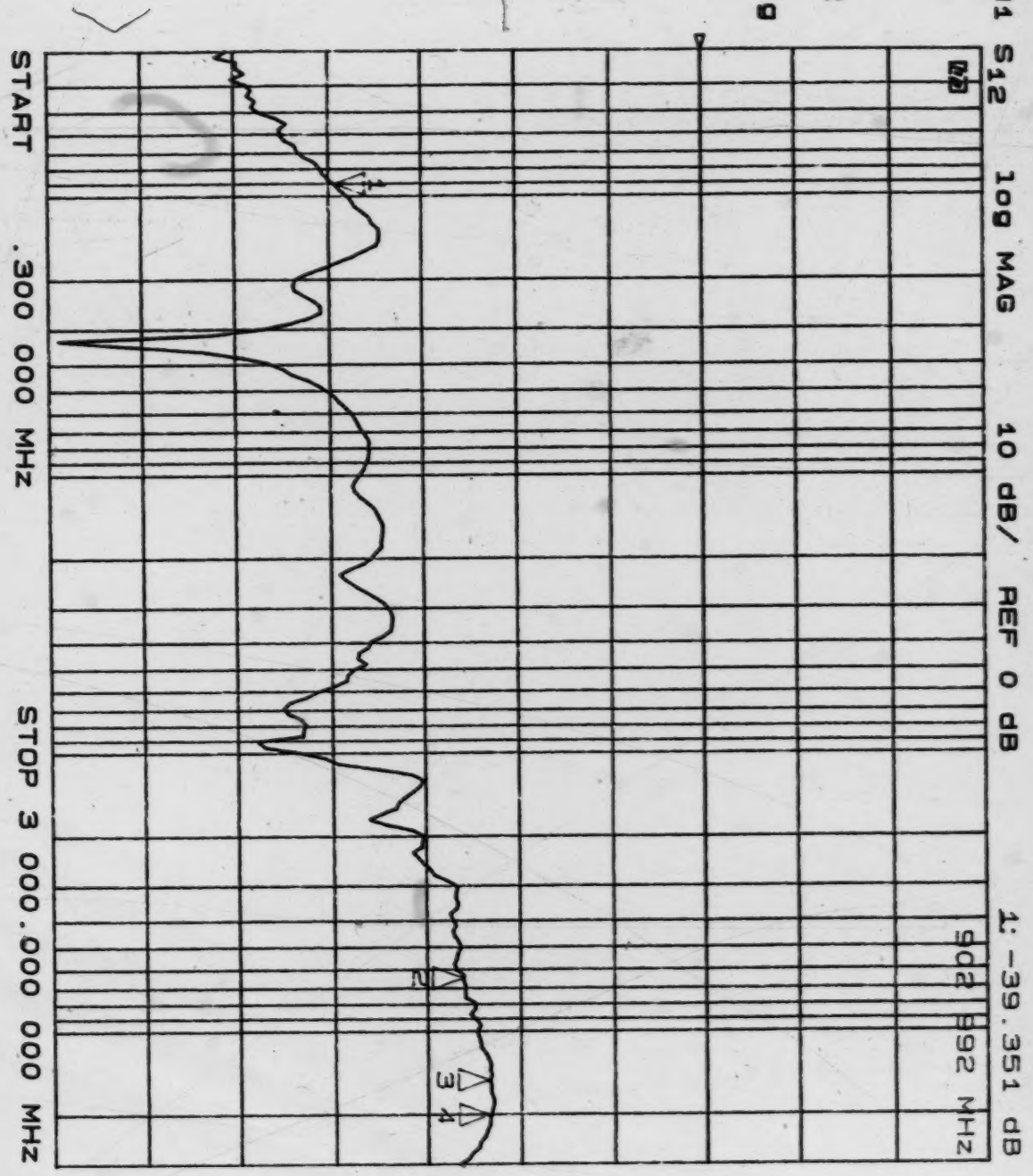


Figure 6.6: S_{12} vs. frequency for LNA1.

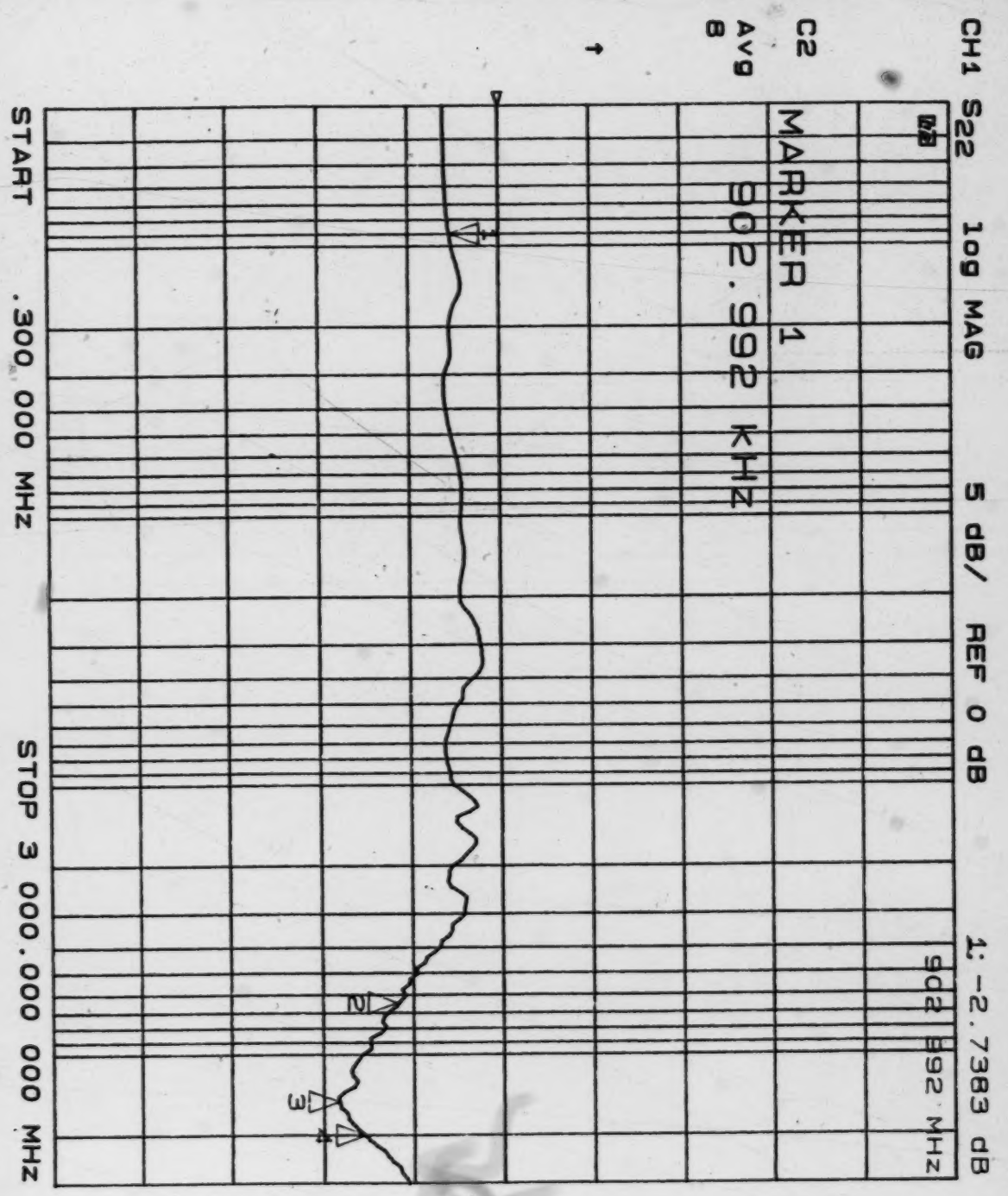


Figure 6.7: S_{22} vs. frequency for LNA1.

Chapter 7

Conclusions

In this chapter, we will summarize the results of the thesis and discuss some related future research topics.

7.1 Thesis Summary

This work attempted to design and build a prototype RF CMOS wideband low-noise amplifier that could be applied in a wireless RF receiver front-end. Several LNA architectures were reviewed, and the methods for wideband amplification were discussed in Chapter 2 of the thesis. Two full-CMOS RF wideband LNAs were designed in a 0.35μm CMOS technology: LNA1 had a 17 dB voltage gain with -3 dB bandwidth of 1.9 GHz, and LNA2 had a 10 dB voltage gain with -3 dB bandwidth of 2 GHz in simulations.

Both of the LNAs were implemented using a bridged T-Coil at the output stage which have been realized by a pair of magnetically coupled inductors, which forms a transformer. This is a new technique for bandwidth extension. The cascoded input stage was used in both designs, so as to minimize the Miller effect, as well as to achieve better gain and higher reverse isolation between the output and the input. The on-chip spiral inductors and transformers were analyzed and simulated using a custom CAD simulator ASITIC (Analysis and Simulation of Inductors and Transformers in Integrated Circuits). The inductors, transformers and the two final LNA circuits were all laid out using 0.35μm CMOS process.

Two LNA circuits were fabricated by CMC, and tested in TRLabs of Edmonton. The measurement results for the first kind of wideband low-noise amplifier, LNA1,

were able to verify the results obtained from the simulations. While LNA2 could not meet the design requirements. The reason was described in Chapter 5 of the thesis.

7.2 Improvements and Future Work

Further extensions and improvements along the lines of the thesis are possible; however, due to limitations of time and scope, these are not pursued in the thesis, but are discussed below:

- Due to the limitation of the time, noise figure was not simulated during the LNA schematic design process. Such a simulation would be done using the Cadence SpectreS simulator.

- Since some instruments for the noise figure measurement were not available at the time of the testing, the measurement of the noise figure was not finished. It would be completed with proper instruments.

- Some corrections for the layout errors discovered in LNA1 and LNA2 will be made, such as widen the power nets which are connected to the power supply pads, and adding the missing load capacitors.

- In addition to the simulation of the schematic, the extracted circuit generated from the layout would be simulated as well. Any necessary changes to the layout would be made before the final design is sent to manufacturing. By this way, more accurate simulation results would be obtained. Furthermore, time and cost will be saved.

- One of the de-embedding techniques (OSD) has been applied in the design; it has not been performed in the testing process. In [51], the detailed analysis on the OSD de-embedding method was described. Several steps can be followed to realize the de-embedding: (1) subtracting the estimated impedance, Z_i, from the measured input and output impedances of the DUT including pads and interconnects; (2) converting the resulting Z-parameters into Y-parameters; and (3) subtracting the estimated admittance, Y_p, from the measured input

and output admittances of the DUT including pads and interconnects. In this case, we assume that

$$\begin{aligned} Z_i &= Z_{i,s}, \\ Y_p &= \frac{1}{Z_{i,o} - Z_i}, \end{aligned}$$

where $Z_{i,s}$ and $Z_{i,o}$ denote the measured input impedance of the short standard and input impedance of the open standard, respectively. A software program written by Mr. Yasser Jamani from CMC will be required in steps 1 and 3.

Bibliography

[1] D. Olcheski, "Semiconductors and related wireless communications activities in Canada," ICTM Branch of Industry Canada, 1998.

[2] B. Razavi, *RF Electronics*, NJ: Prentice Hall PTR, 1997.

[3] L. E. Larson, *RF and Microwave Circuit Design for Wireless Communications*, Norwood, MA: Artech House, 1996.

[4] Y. E. Papanaos, *Radio-Frequency Microelectronic Circuits for Telecommunication Applications*, Norwell, MA: Kluwer Academic Publishers, 1999.

[5] W. M. Leach, "Fundamentals of low-noise analog circuit design," Proc. IEEE, vol. 82, no. 10, pp.1515-1538, Oct. 1994.

[6] P. R. Gray and P. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, NY: Wiley, 1993.

[7] R. Sagers, "Intercept point and undesired responses," Trans. IEEE Veh. Technol., vol. VT-32, no. 1, pp. 121-133, Feb. 1983.

[8] A. R. Hambley, *Electronics: A Top-Down Approach to Computer-Aided Circuit Design*, NJ: Prentice Hall, 1994.

[9] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, NY: Cambridge University Press, 1998.

[10] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," IEEE J. of Solid-Sate Circuits, vol. 32, no. 5, pp. 745-759, May 1997.

[11] A. Van der Ziel, "Noise in solid-state devices and lasers," Proc. IEEE, vol. 58, Aug. 1970.

[12] Y. Ge and K. Mayaram, "A comparative analysis of CMOS low noise amplifier for RF applications," Proc. IEEE Int. Symp. Circuits and Systems, June 1998.

[13] L. M. Burns, "Applications for GaAs and silicon integrated circuits in next generation wireless communication systems," IEEE J. of Solid-State Circuits, vol. 30, no. 10, pp. 1088-1095, Oct. 1995.

[14] K. B. Ashby, I. A. Koullias, W. C. Finley, J. J. Basek, and S. Moinian, "High Q inductors for wireless applications in a complementary silicon bipolar process," IEEE J. Solid-State Circuits, vol. 31, no. 1, pp. 4-9, Jan. 1996.

[15] H. M. Greenhouse, "Design of planar rectangular microelectronic inductors," IEEE Transactions on Parts, Hybrids and Packaging, vol. PHP-10, no. 2, pp. 101-109, June 1974.

[16] N. M. Nguyen and R. G. Meyer, "Si IC-compatible inductors and LC passive filters," IEEE J. of Solid-State Circuits, vol. 25, no. 4, pp. 1028- 1031, Aug. 1990.

[17] J. J. Zhou, and D. J. Allstot, "Monolithic transformers and their applications in differential CMOS RF low-noise amplifier," IEEE J. of Solid-State Circuits, vol. 33, no. 12, pp. 2020-2027, Dec. 1998.

[18] S. S. Mohan, C. P. Yue, M. D. Hershenson, S. S. Wong, and T. H. Lee, "Modeling and characterization of on-chip transformers," IEDM 1998, pp. 19.3.1-19.3.4.

[19] A. M. Niknejad, "Analysis, design, and optimization of spiral inductors and transformers for Si RF ICs," Master of Philosophy thesis, University of California, Berkeley, 1998.

[20] J. R. Long and M. A. Copeland, "The modeling, characterization, and design of monolithic inductors for silicon RF IC's," IEEE J. of Solid-State Circuits, vol. 32, no. 3, pp. 357-359, March 1997.

[21] C. W. Davidson, *Wideband Voltage Amplifiers*, International Textbook Company Limited, 1974.

[22] J. D. Lenk, *Simplified Design of IC Amplifiers*, Butterworth-Heinemann, 1996.

[23] C. H. Evans, *Electronic Amplifiers*, NY: Litton Education Publishing, Inc., 1979.

[24] S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, McGraw-Hill, Inc., 1988.

[25] J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large suspended inductors on silicon and their use in a 2-μm CMOS RF amplifier," IEEE Electron Device Letters, vol. 14, no. 5, pp. 246-248, May 1993.

[26] A. N. Karanicolas, "A 2.7V 900MHz CMOS LNA and mixer," IEEE J. of Solid-State Circuits, vol.31, no. 12, pp. 1939-1944, Dec. 1996.

[27] A. Rofougaran, J. Y.-C. Chang, M. Rofougaran, and A. A. Abidi, "A 1GHz CMOS RF front-end IC for a direct-conversion wireless receiver," IEEE J. of Solid-State Circuits, vol. 31, no. 7, pp. 880-889, July 1996.

[28] S. Sheng *et al.*, "A low-power CMOS chipset for spread-spectrum communications," in ISSCC Dig. Tech. Papers, vol. 39, pp. 346-347, 1996.

[29] Q. Huang, F. Piazza, P. Orsatti, and T. Ohguro, "The impact of scaling down to deep-submicron on CMOS RF circuits," IEEE J. of Solid-State Circuits, vol. 33, pp. 1023-1036, July 1998.

[30] Q. Huang, P. Orsatti, and F. Piazza, "GSM transceiver front-end circuits in 0.25-μm CMOS," IEEE J. of Solid-State Circuits, vol. 34, no. 3, pp. 292-302, Mar. 1999.

[31] C. Y. Wu and S. Y. Hsiao, "The design of a 3-V 900-MHz CMOS bandpass amplifier," IEEE J. of Solid-State Circuits, vol. 32, no. 2, pp. 159-168, Feb. 1997.

[32] Alan B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, John Wiley & Sons, Inc., 1984.

[33] Dennis L. Feucht, *Handbook of Analog Circuit Design*, Academic Press, Inc., 1990.

[34] Troels E. Kolding, "Review of RF CMOS performance and future process innovations," Technical Report R98-1014, Oct. 1998.

[35] J. Crols and M. S. J. Steyaert, *CMOS Wireless Transceiver Design*, Kluwer Academic Publishers, 1997.

[36] L. E. Larson, "Radio frequency integrated circuit technology for low-power wireless communications," IEEE Personal Communications, vol. 5, no. 6, pp. 11-19, June 1998.

[37] S. Sheng and R. Brodersen, *Low-Power Wireless Communications: A Wideband CDMA System Design*, Kluwer Academic Publishers, 1998.

[38] M. L. McMahan, A. Khatzibadeh, and P. Shah, *Wireless Systems and Technology Overview*, Texas Instruments, Inc., White paper, 1997.

[39] C. Wann, F. Assaderaghi, L. Shi, K. Chan, S. Cohen, H. Hovel, K. Jenkins, Y. Lee, D. Sadana, R. Viswanathan, S. Wind, and Y. Taur, "High-performance 0.07-μm CMOS with 9.5-ps gate delay and 150 GHz f_T," IEEE Electron Device Letters, vol. 18, no. 12, pp. 625-627, Dec. 1997.

[40] R. Jacob Baker, Harry W. Li, and David E. Boyce, *CMOS Circuit Design, Layout, and Simulation*, Institute of Electrical and Electronics Engineering, Inc., 1998.

[41] Henrik Sjöland, *Highly Linear Integrated Wideband Amplifiers: Design and Analysis Techniques for Frequencies from Audio to RF*, Kluwer Academic Publishers, 1999.

[42] Arvin R. Shahani, Derek K. Shaeffer, and Thomas H. Lee, "A 12-mW wide dynamic range CMOS front-end for a portable GPS receiver," IEEE J. of Solid-State Circuit, vol. 32, no. 12, pp. 2061-2070, Dec. 1997.

[43] Qiuting Huang, Paolo Orsatti, Francesco Piazza, and T. Yoshitomi, "Broadband, 0.25μm CMOS LNAs with sub-2dB NF for GSM applications," IEEE 1998 Custom Integrated Circuits Conference, pp. 67-70, 1998.

[44] S. M. Sze, *Physics of semiconductor devices*, New York: Wiley, 1981.

[45] B. Razavi, R. H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," IEEE Trans. Circuits Syst. I, vol. 41, pp. 750-754, Nov. 1994.

[46] D. Johns and K. Martin, *Analog Integrated Circuit Design*, NY: Wiley, 1996.

[47] Eric Strid, "On-wafer measurements with the HP 8510 network analyzer and Cascade microtech wafer probes," RF & Microwave Measurement Symposium and Exhibition.

[48] "HP 8753A Network Analyzer: system verification," Hewlett Packard.

[49] David D. Clegg, "TRLabs Edmonton RF manual," TRLabs, 2000.

[50] M. C. A. M. Koolen, J. A. M. Geelen and M. P. J. G. Versleijen, "An improved de-embedding technique for on-wafer high-frequency characterization," IEEE 1991 Bipolar Circuits and Technology Meeting 8.1, pp. 188-191, 1991.

[51] T. E. Kolding, "On-wafer calibration techniques for Giga-Herz CMOS measurements," Proc. IEEE 1999 Int. Conf. on Microelectronic Test Structures, vol. 12, pp. 105-110, March 1999.

[52] I. M. Filanovsky, "EE671: The Design of CMOS Radio-Frequency Integrated Circuits," Class notes, May 1999.

[53] J. Davidse, *Integration of Analogue Electronic Circuits*, Academic Press Inc. (London) Ltd., 1979.

Appendix A

Appendix

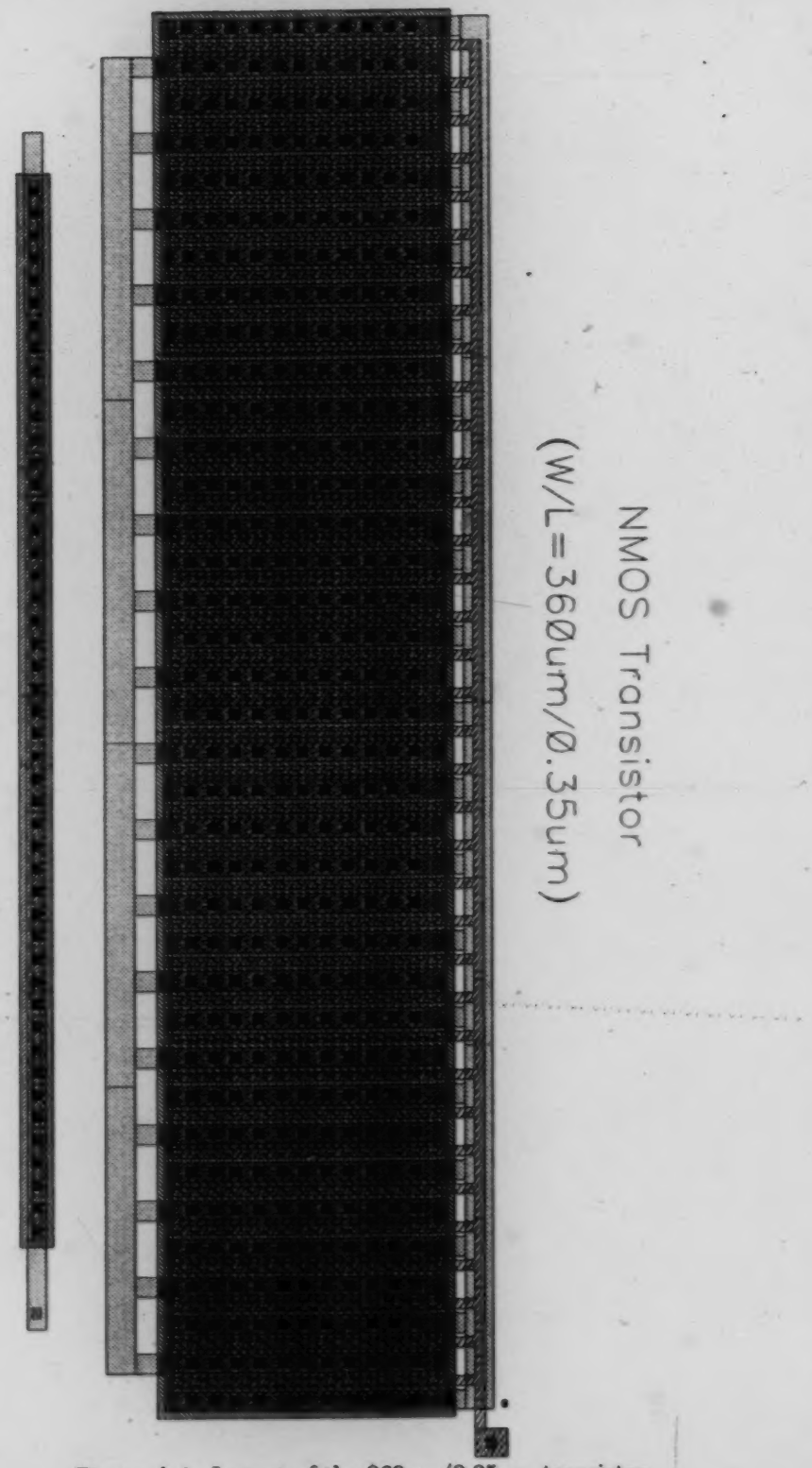


Figure A.1: Layout of the 360μm/0.35μm transistor.

Figure A.2: Layout of the 90fF poly-poly capacitor.

Figure A.3: Layout of the 200Ω poly-silicon resistor.

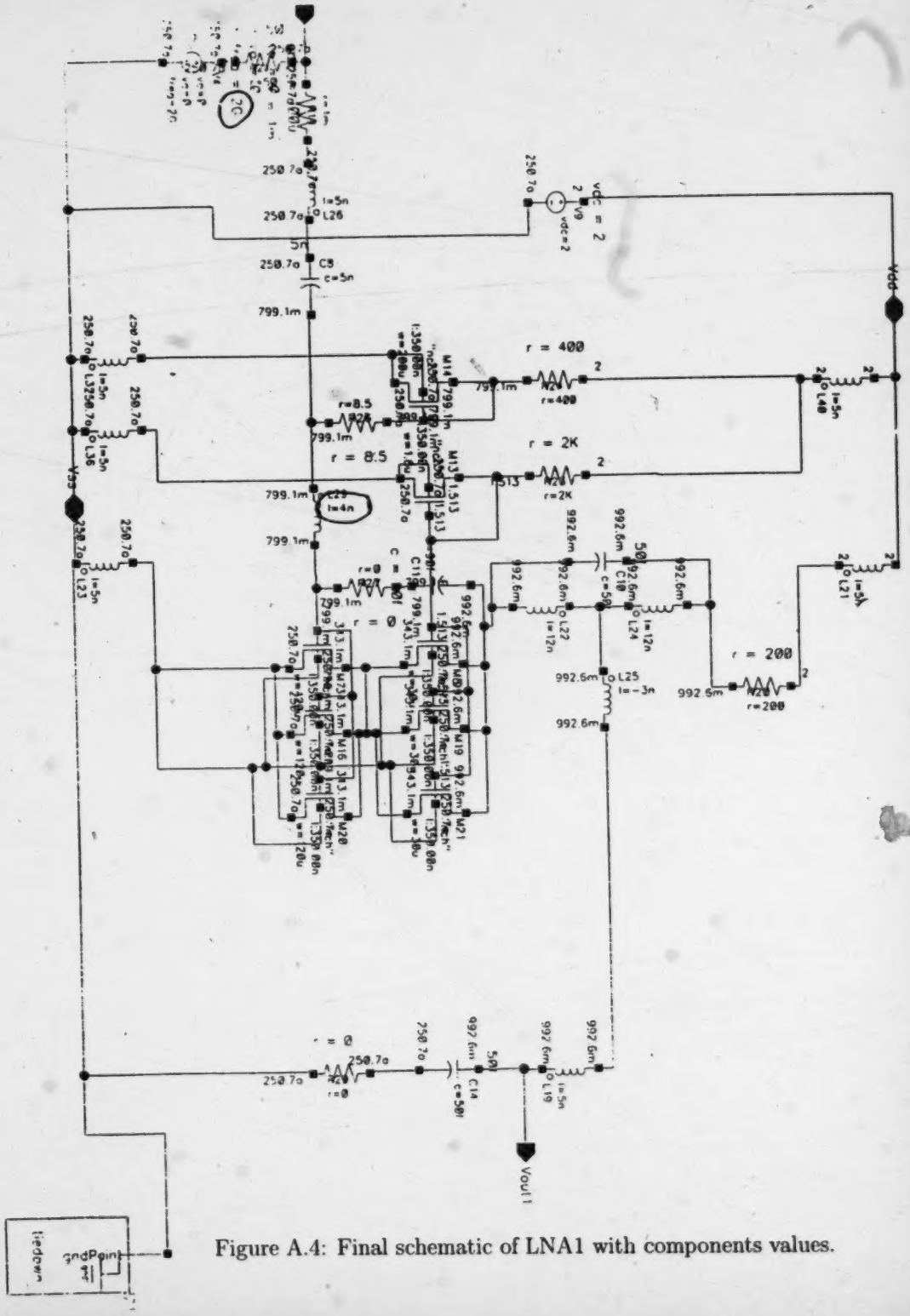


Figure A.4: Final schematic of LNA1 with components values.

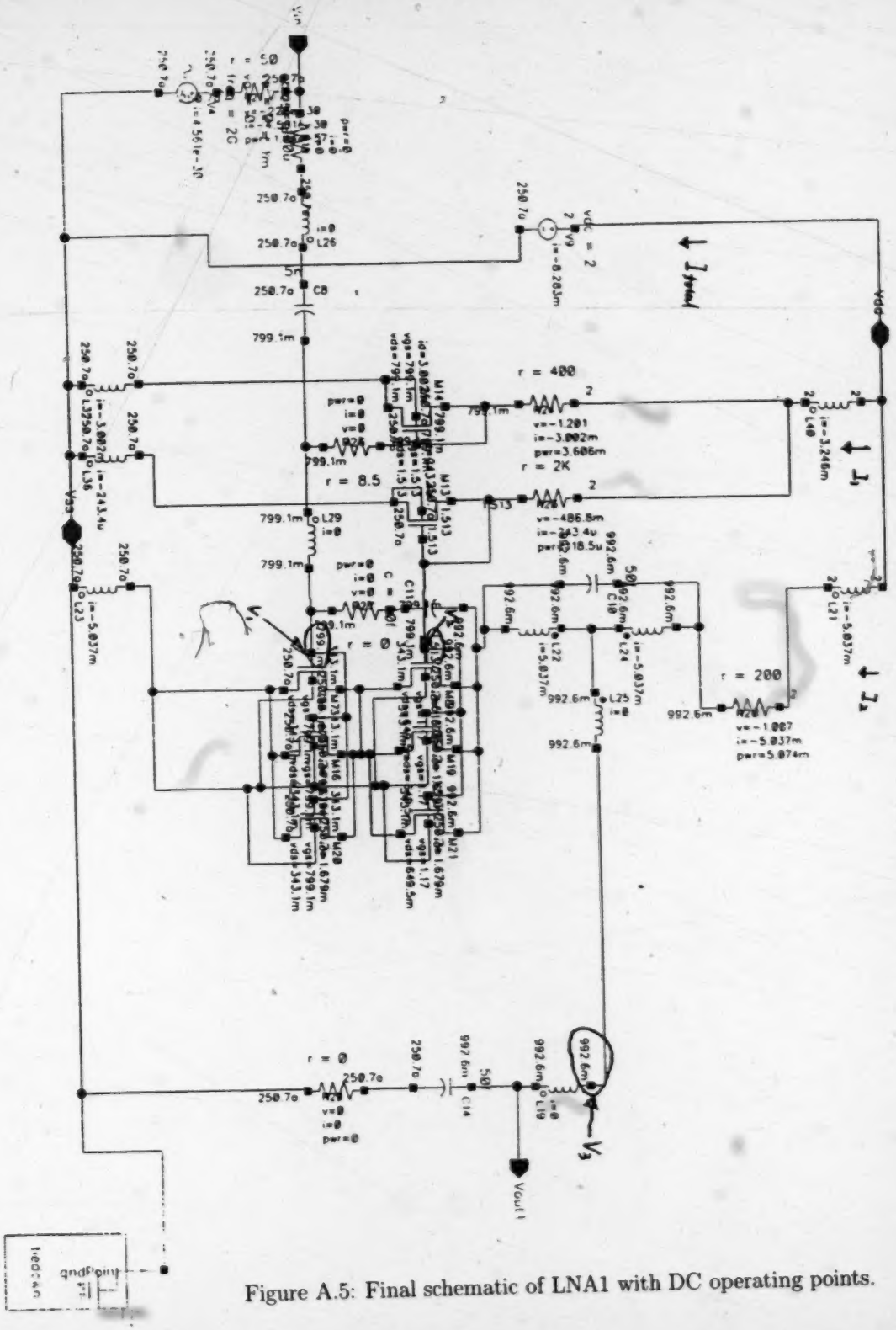


Figure A.5: Final schematic of LNA1 with DC operating points.

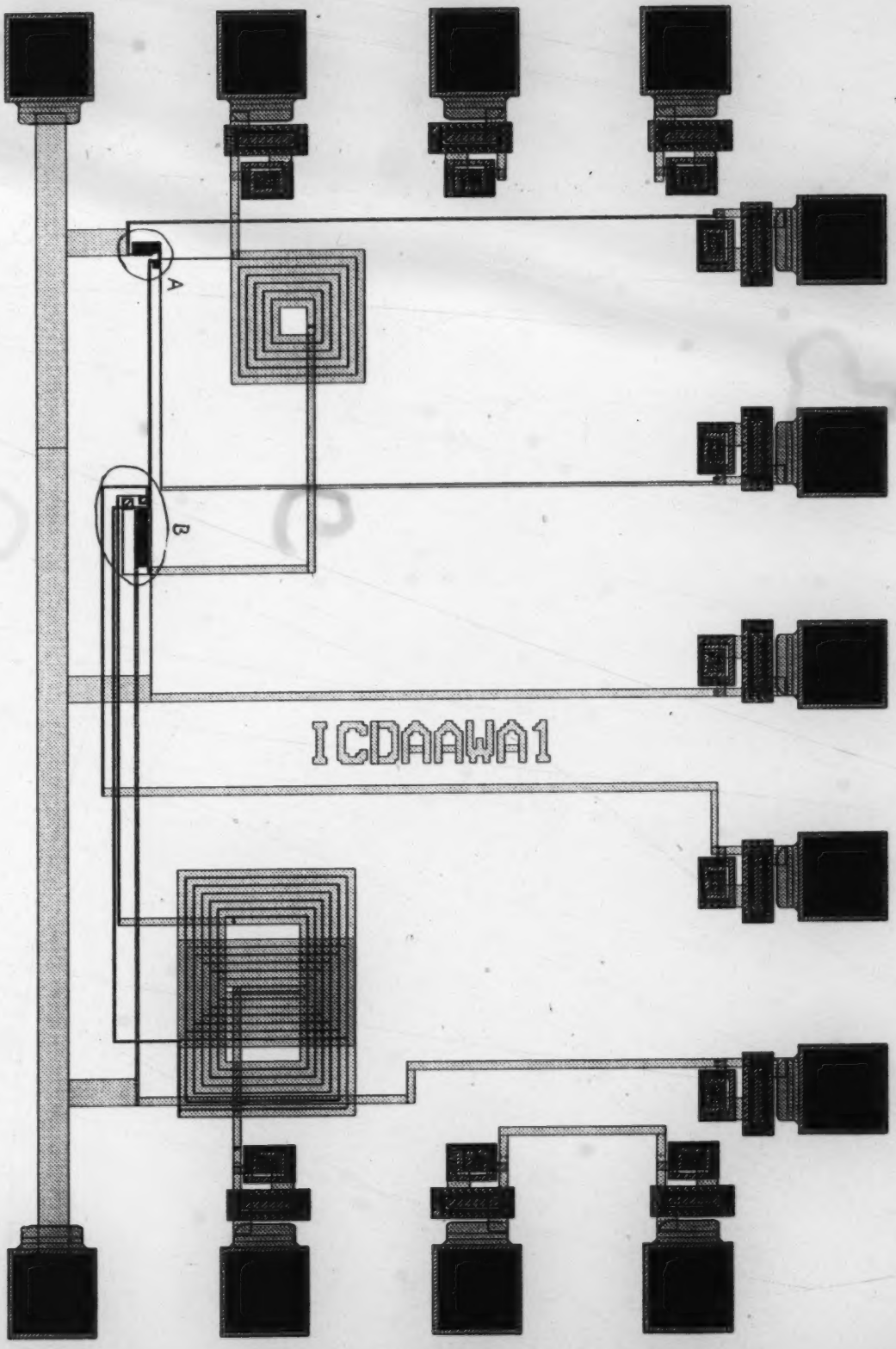


Figure A.6: Final layout of LNA1.

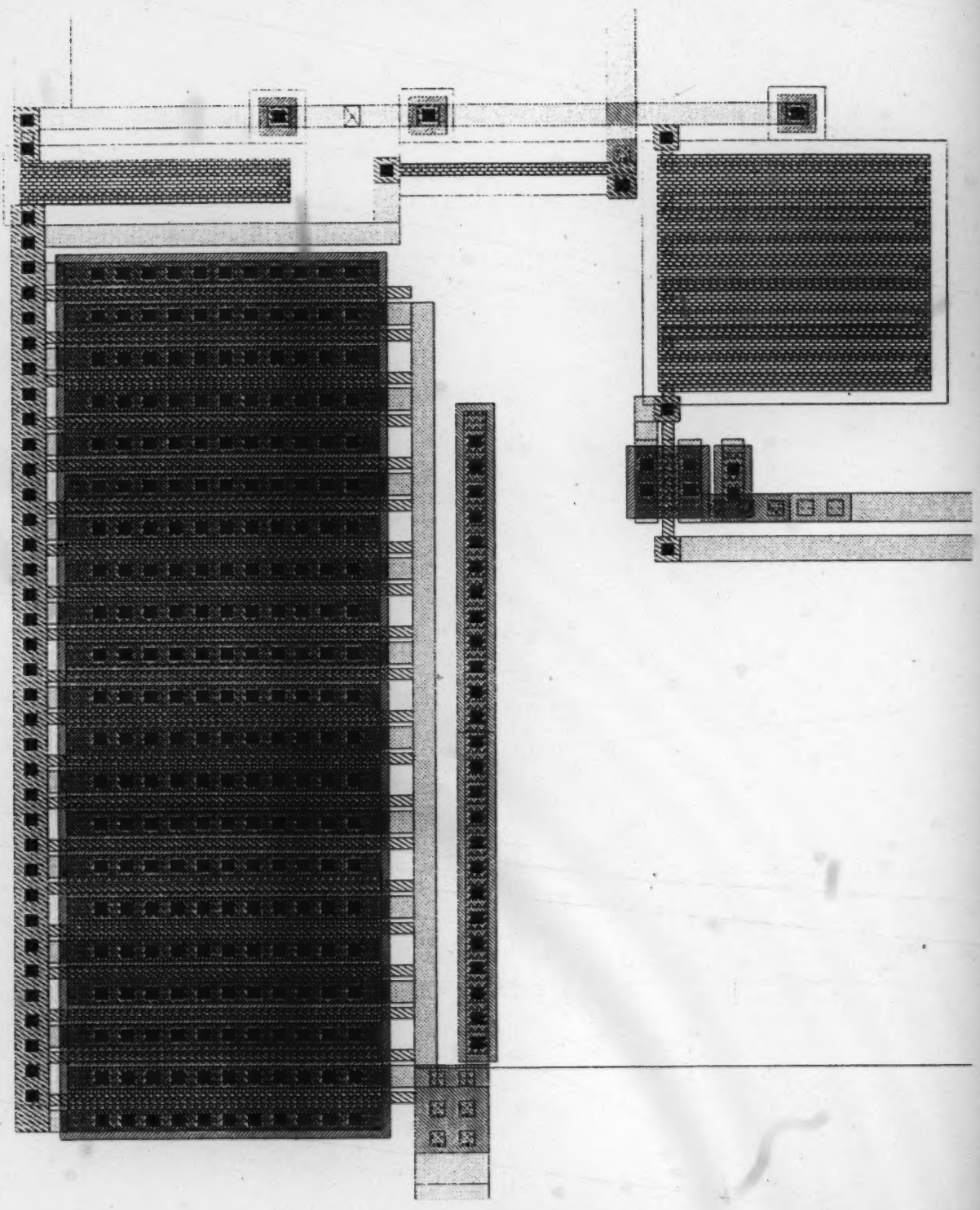


Figure A.7: Part A in Figure A.6 in detail.

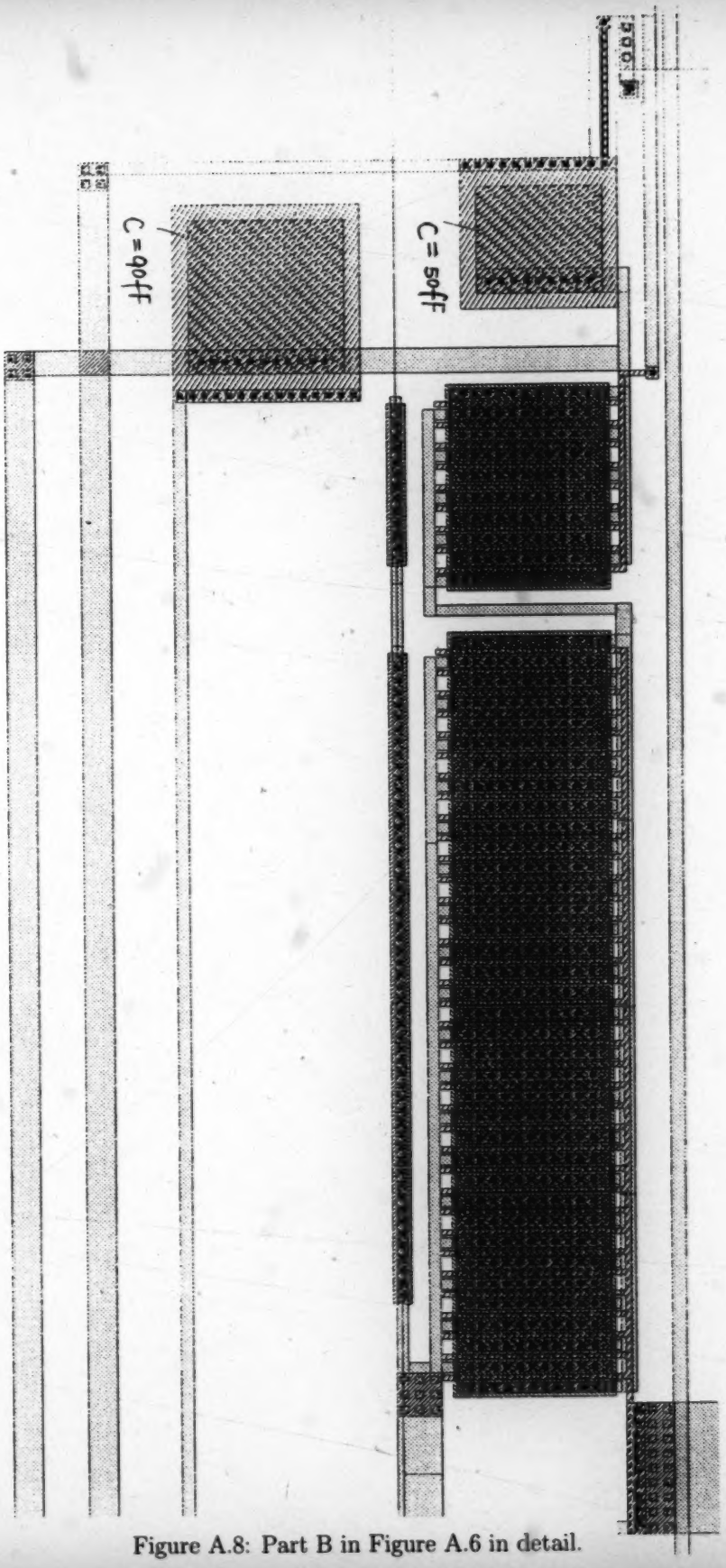


Figure A.8: Part B in Figure A.6 in detail.

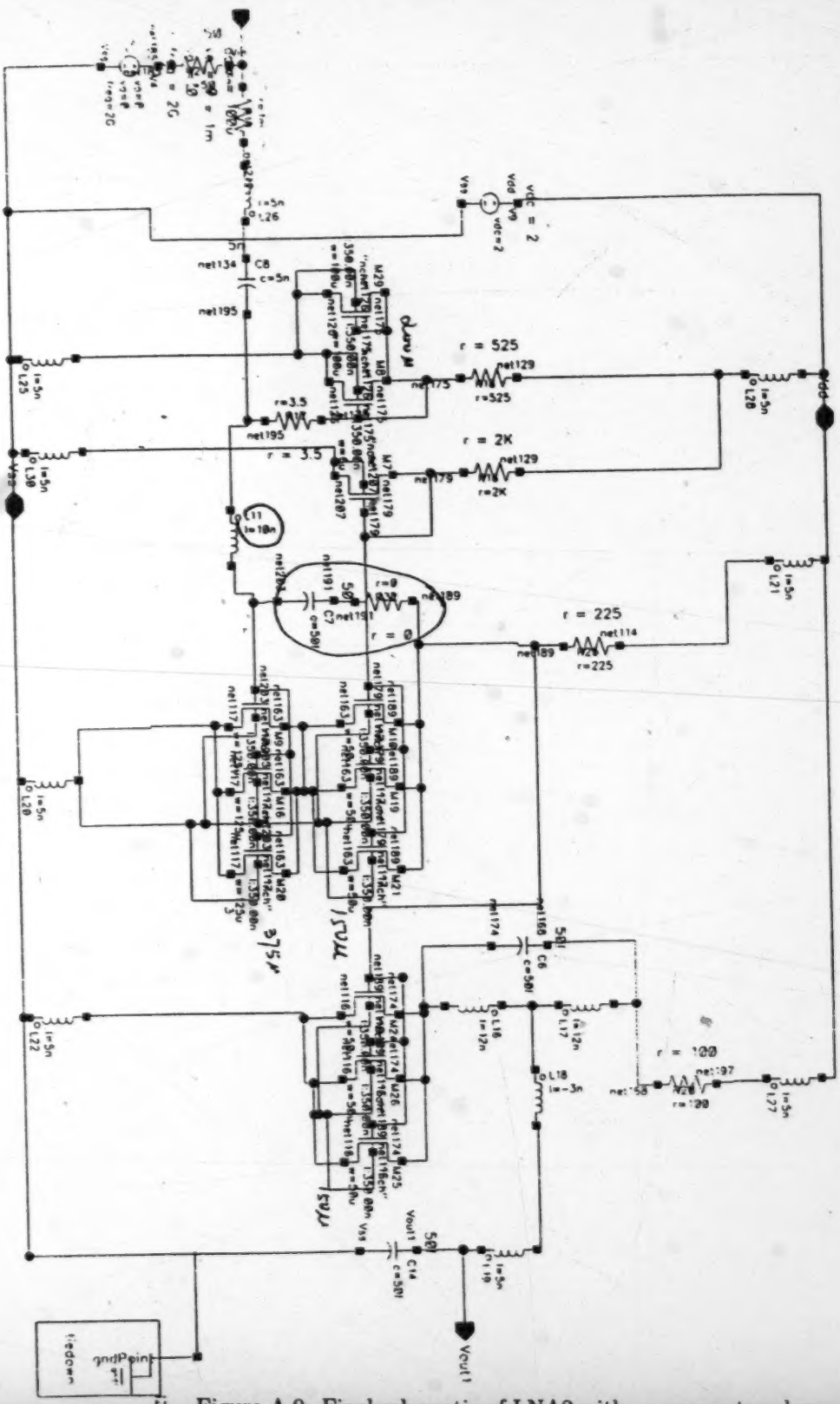


Figure A.9: Final schematic of LNA2 with components values.

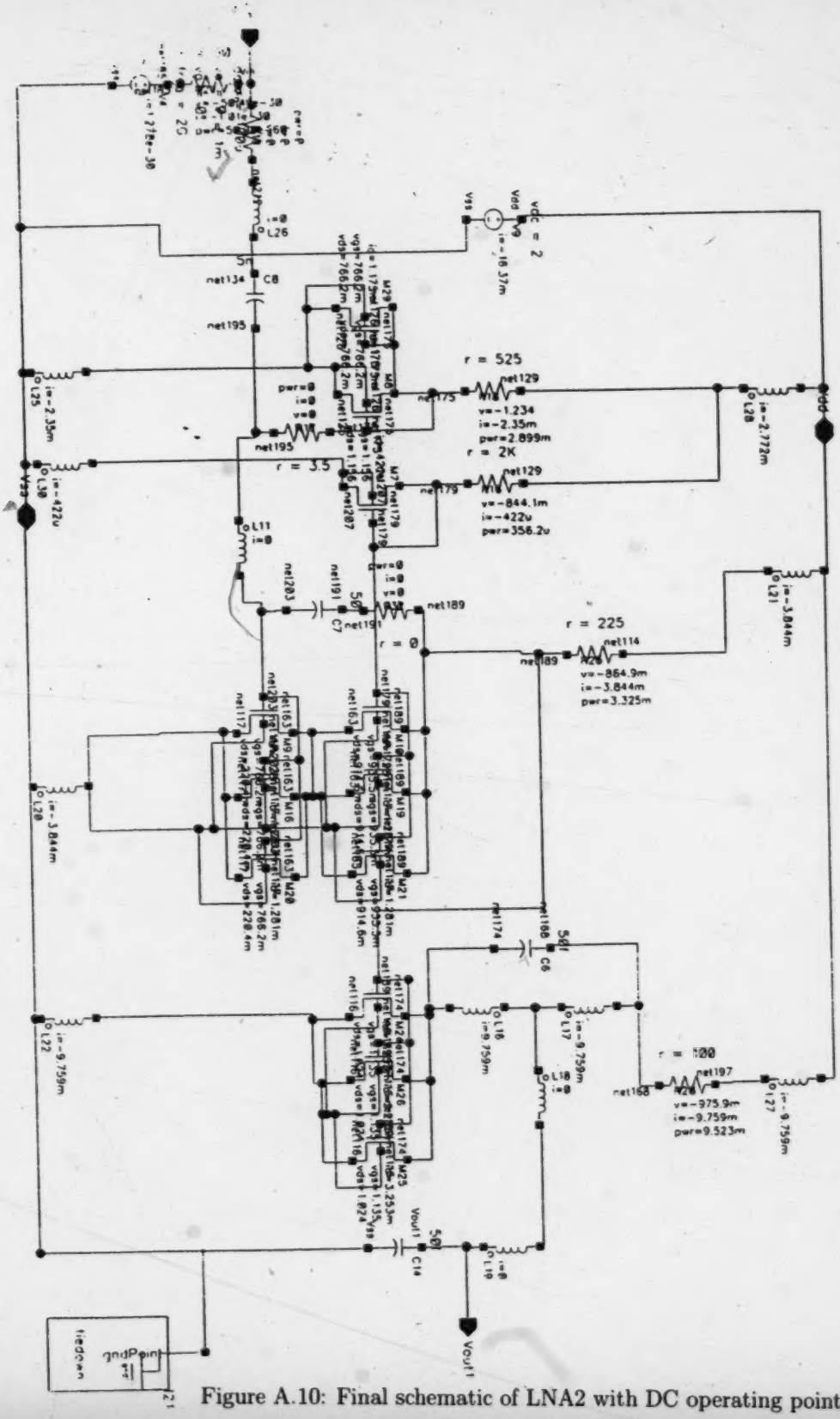


Figure A.10: Final schematic of LNA2 with DC operating points.

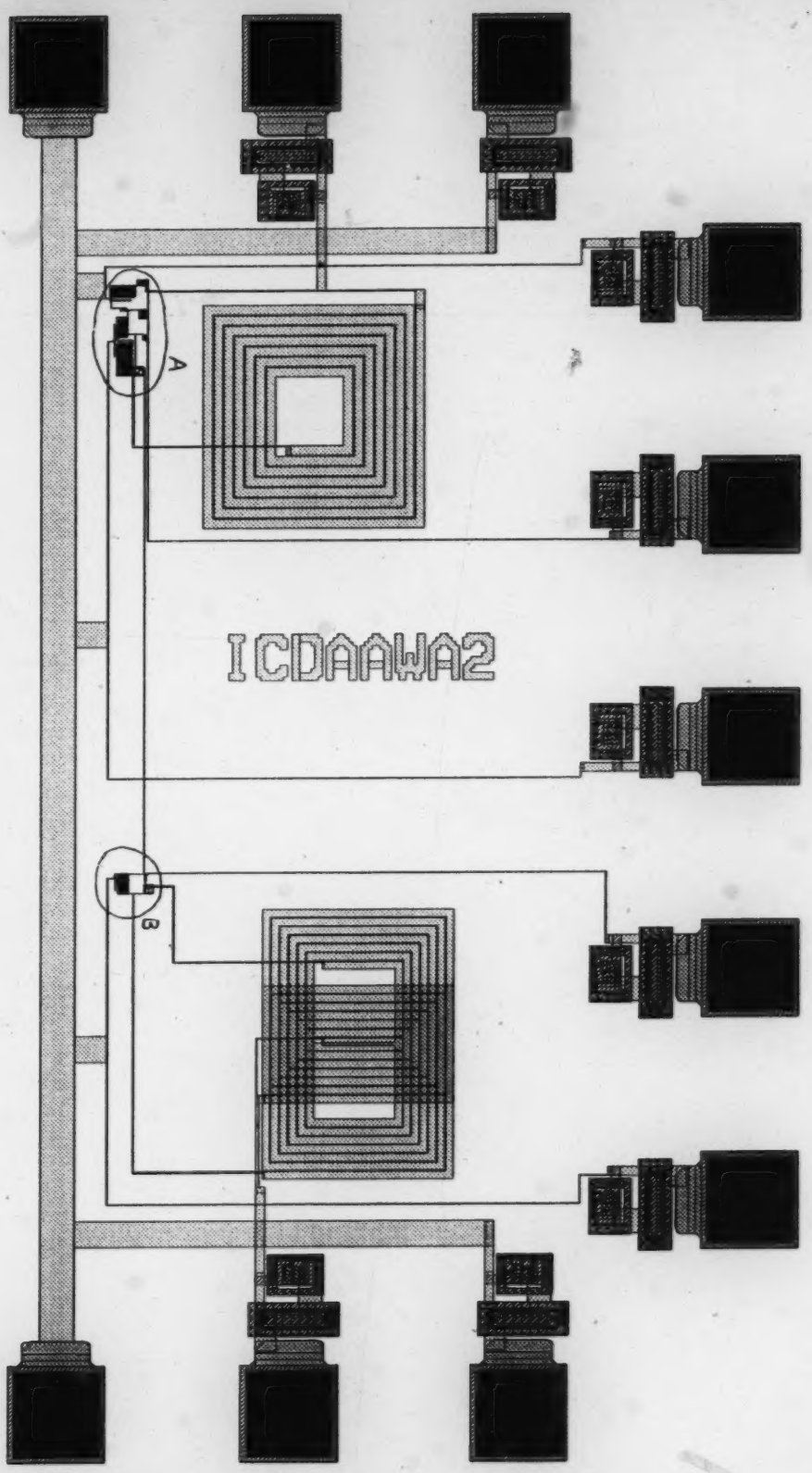


Figure A.11: Final layout of LNA2.

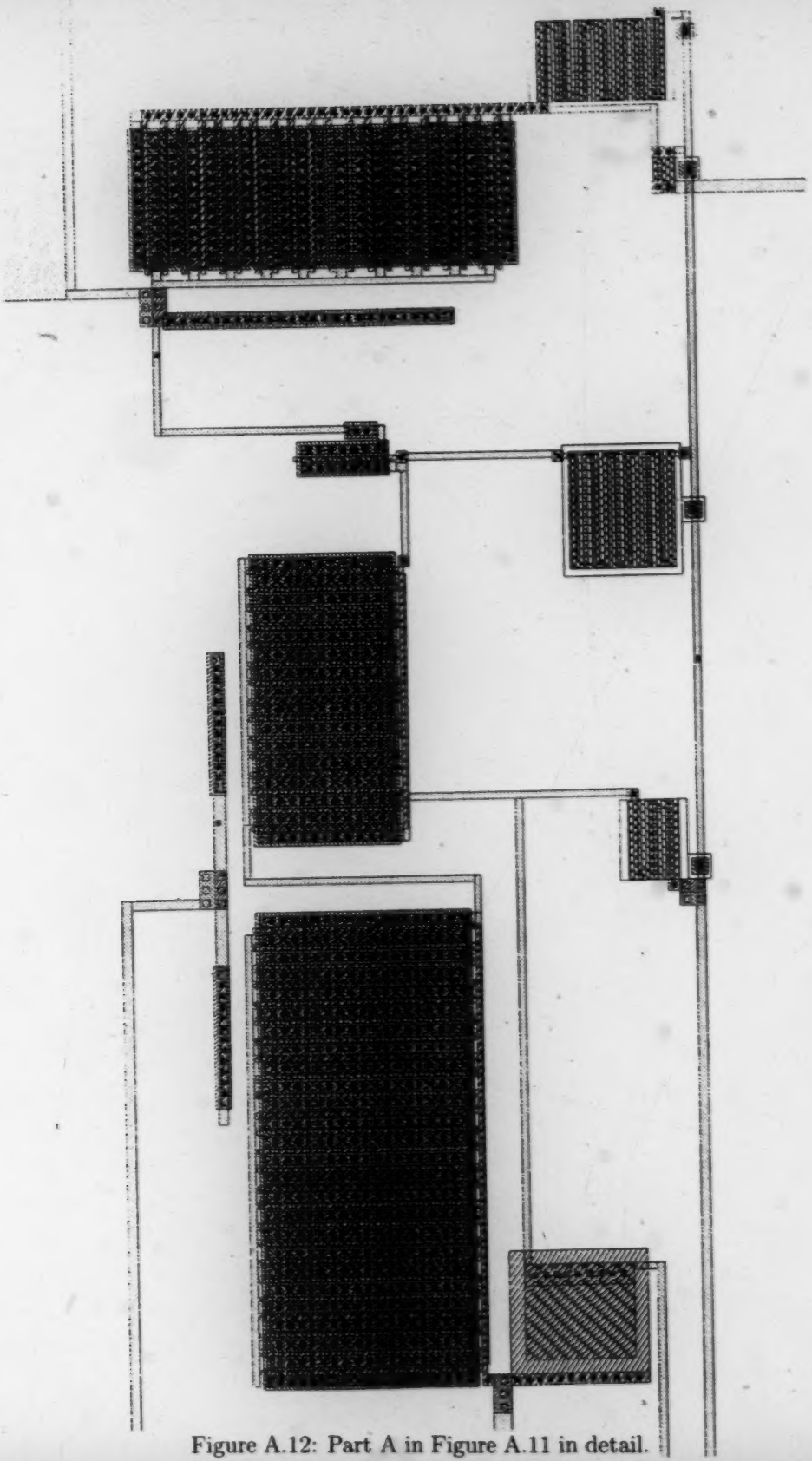


Figure A.12: Part A in Figure A.11 in detail.

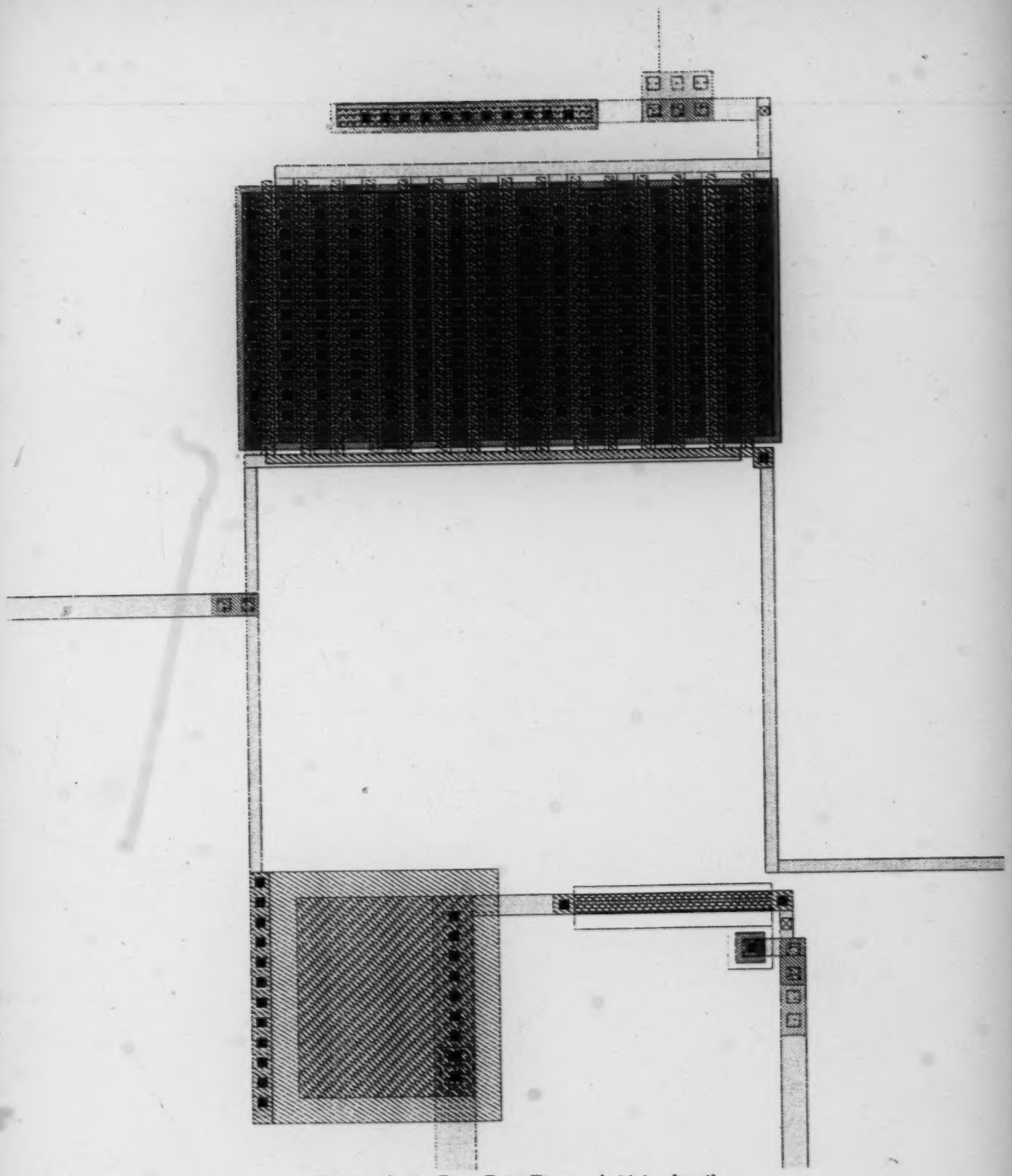


Figure A.13: Part B in Figure A.11 in detail.

Figure A.14: (a) V_{out} vs. frequency for LNA1, (b) Z_{in} vs. frequency for LNA1.

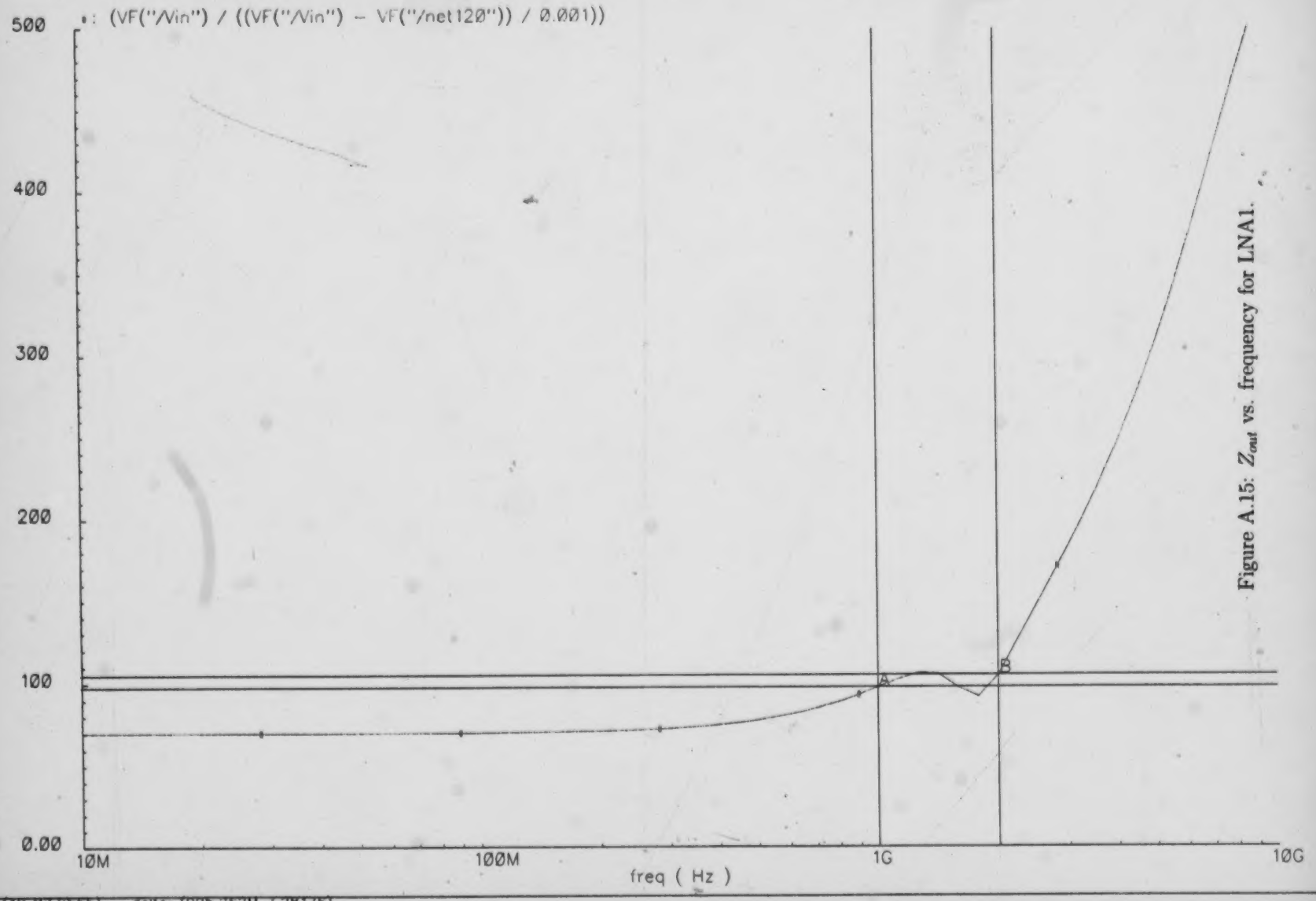


Figure A.15: Z_{out} vs. frequency for LNA1.

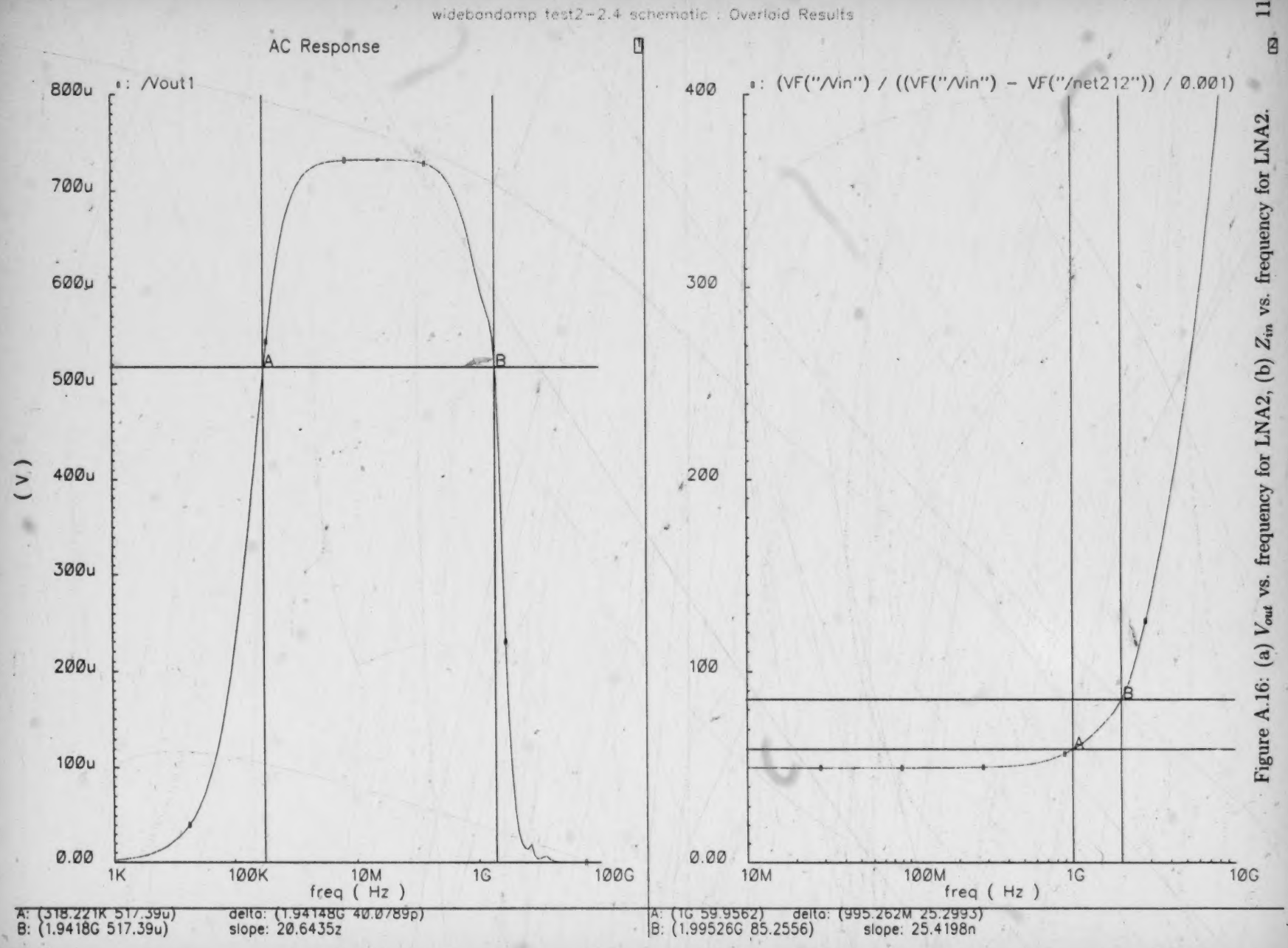


Figure A.16: (a) V_{out} vs. frequency for LNA2, (b) Z_{in} vs. frequency for LNA2.

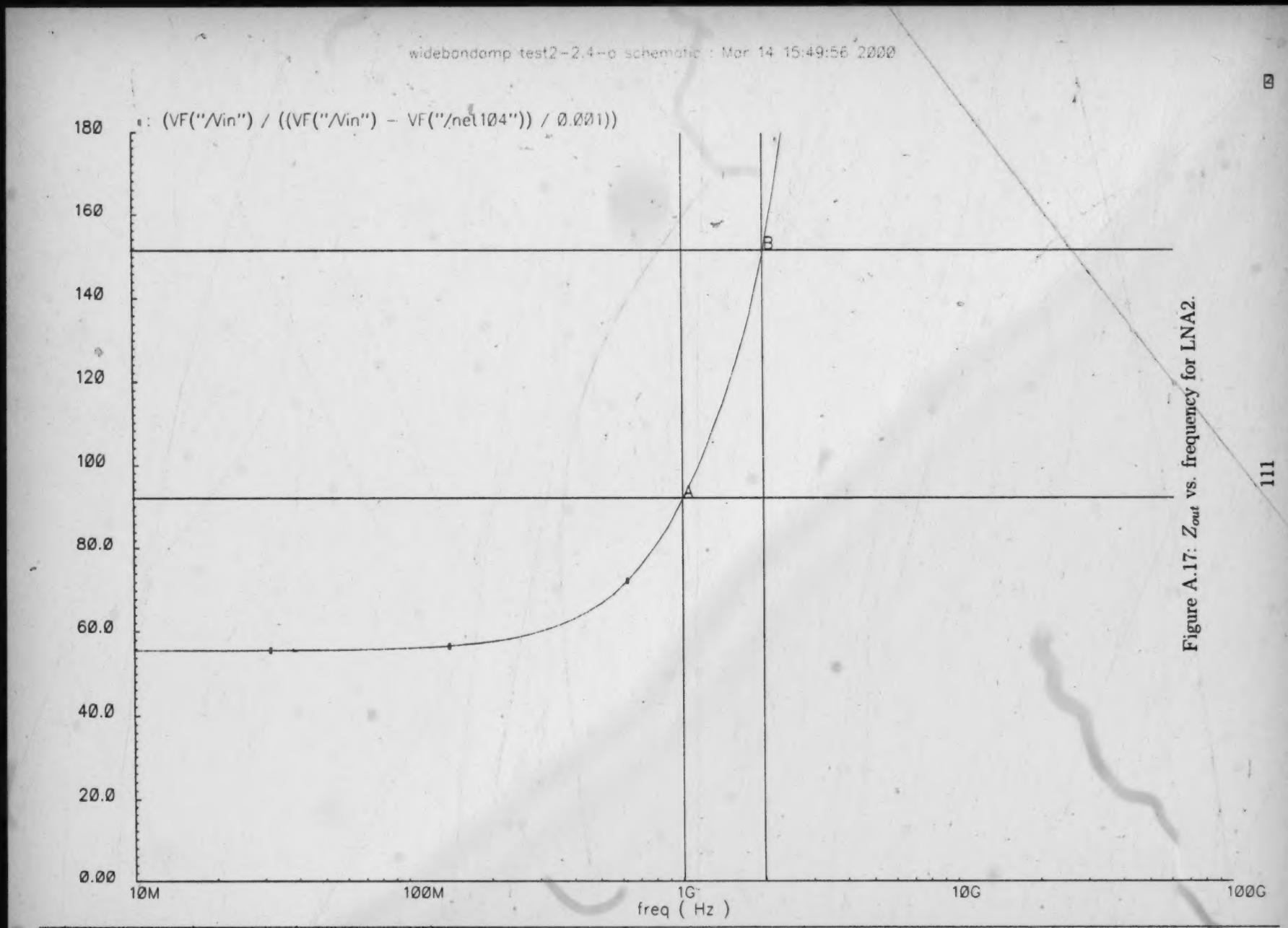


Figure A.17: Z_{out} vs. frequency for LNA2.

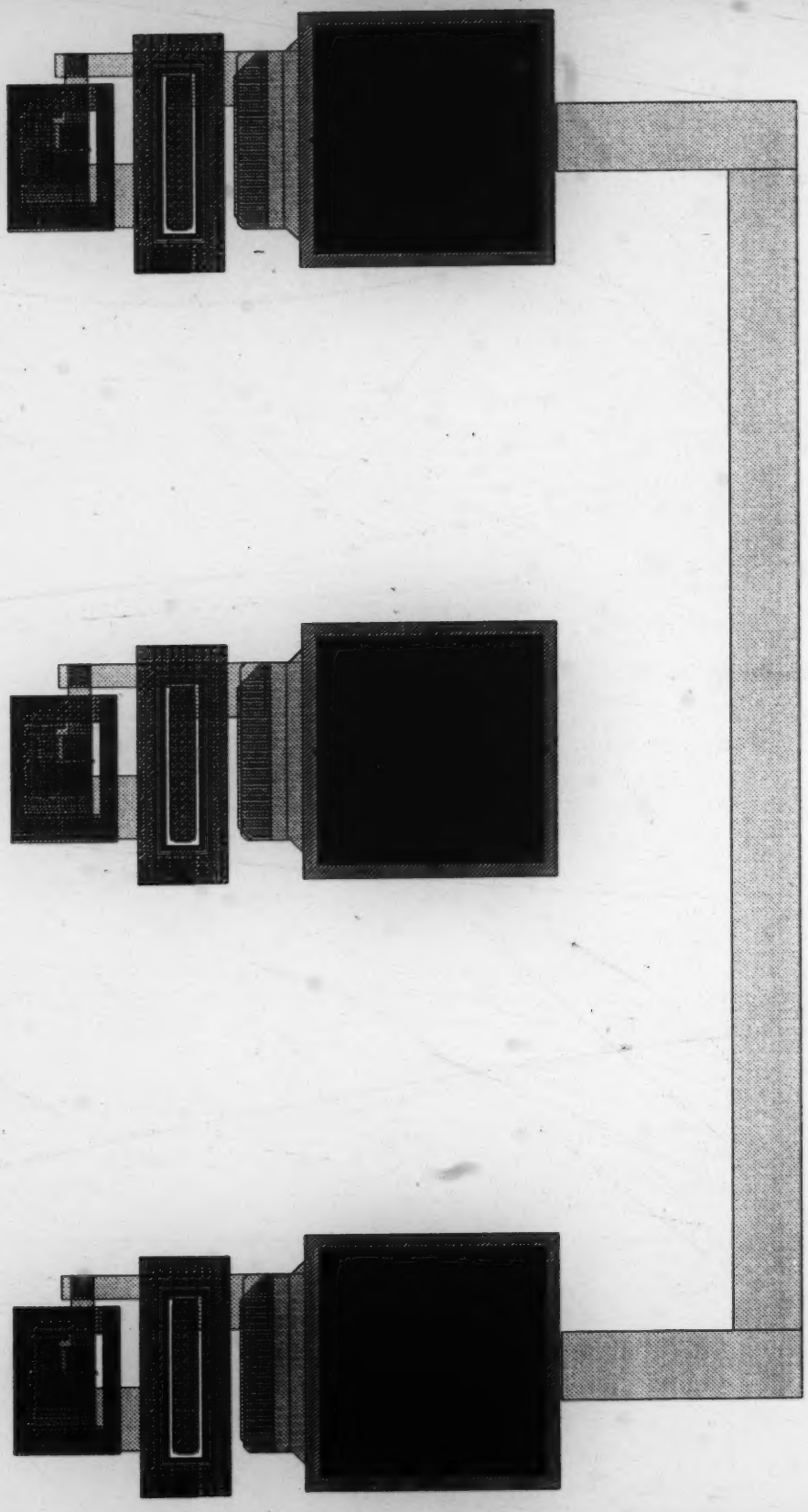


Figure A.18: Pad layout used throughout the project.

Appendix B

Appendix

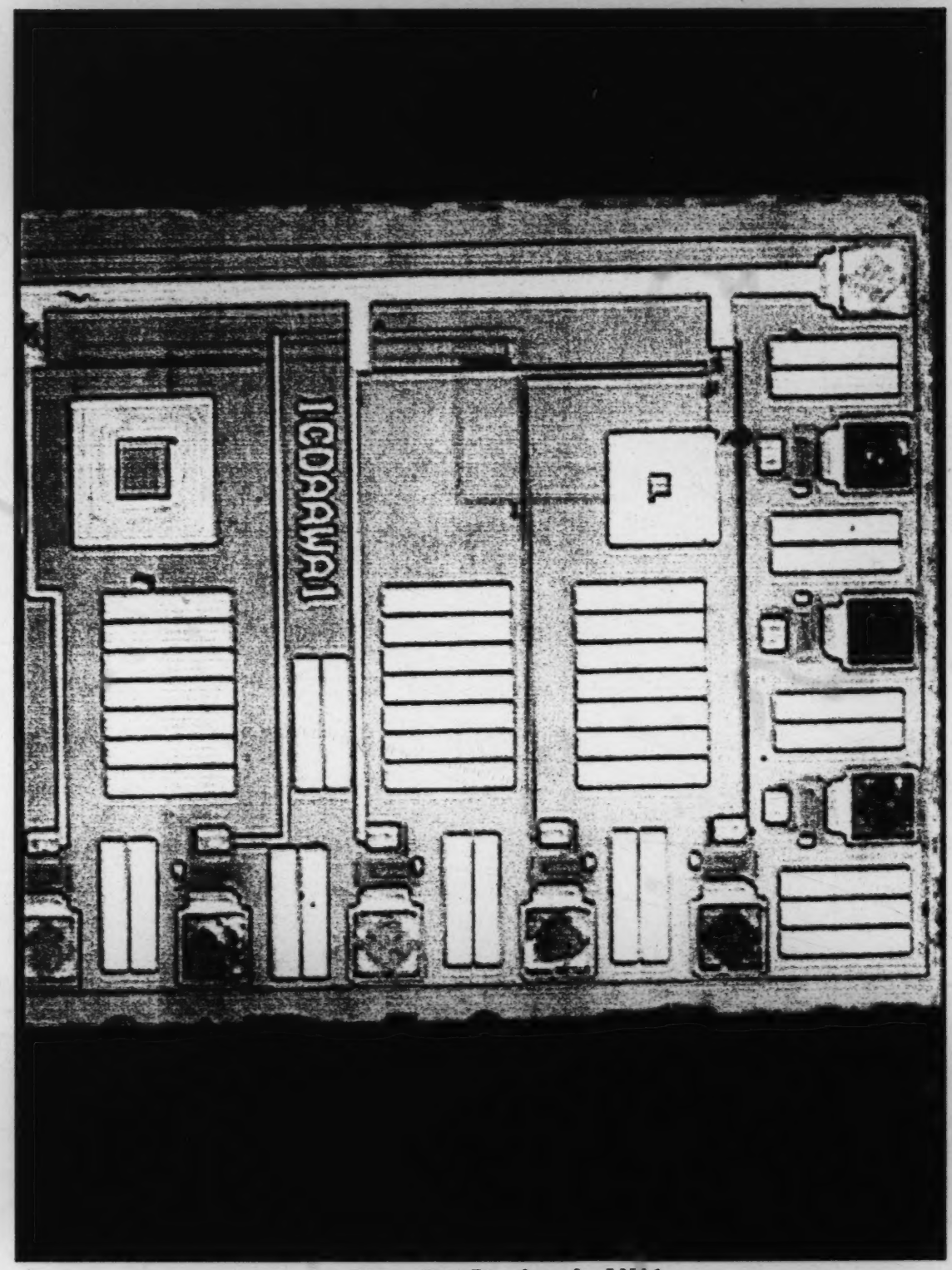


Figure B.1: Die photo for LNA1.

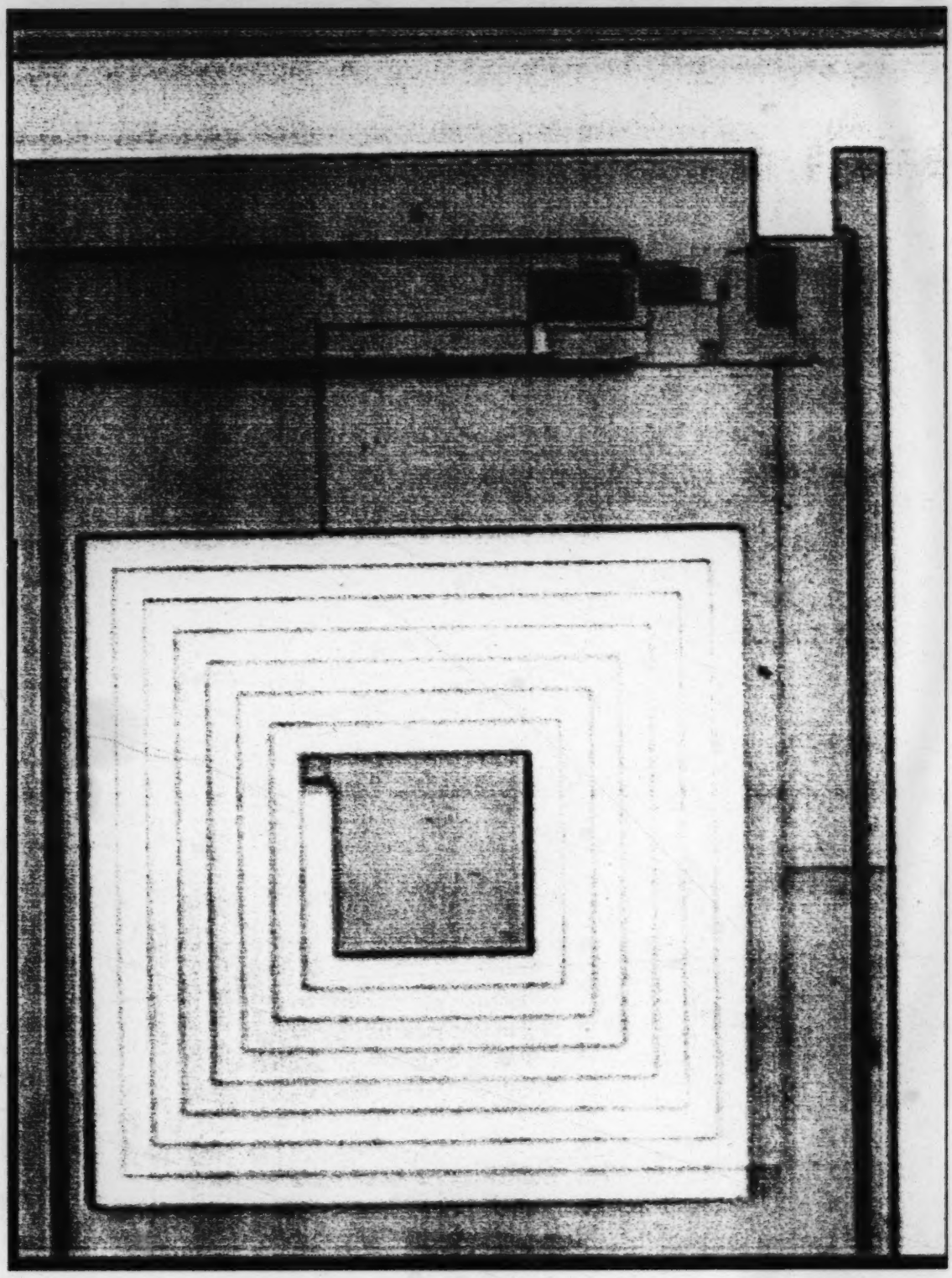


Figure B.2: Die photo for the inductor with inductance of 10nH.

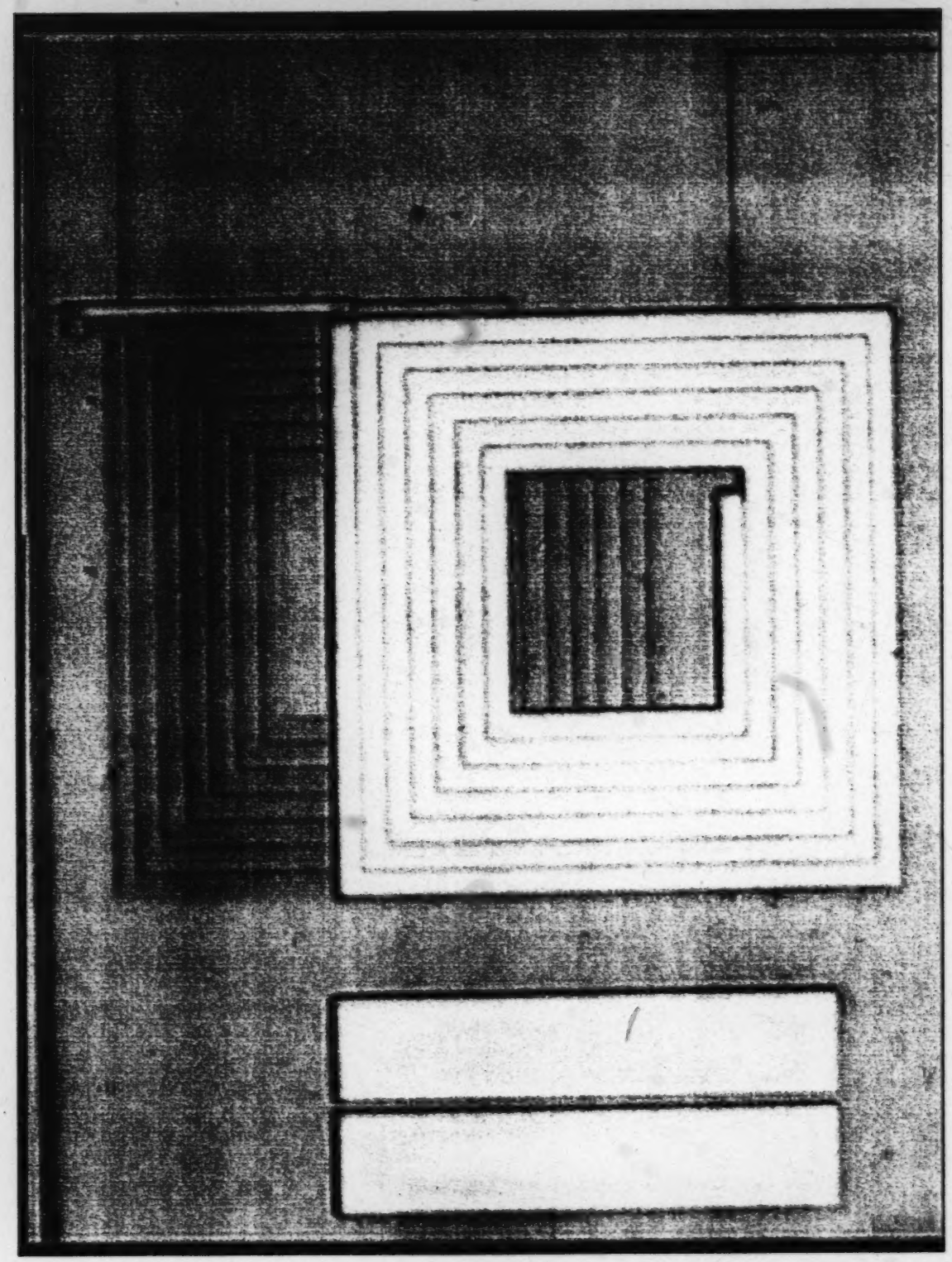


Figure B.3: Die photo for the stacked transformer (P=9nH; S=9nH; k=1/3).